d

**UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE FÍSICA DE SÃO CARLOS
DEPARTAMENTO DE FÍSICA E INFORMÁTICA**

**“ESTRATÉGIA ATENCIONAL PARA BUSCA VISUAL E RECONHECIMENTO
INVARIANTE DE OBJETOS BASEADA NA INTEGRAÇÃO DE
CARACTERÍSTICAS BOTTOM-UP E TOP-DOWN”**

EVELINA MARIA DE ALMEIDA NEVES

Tese apresentada ao Instituto de Física de São Carlos, Universidade de São Paulo, para
obtenção do título de Doutor em Ciências “Física Aplicada-Opção: Física Computacional

Orientador: Prof. Dr. Adilson Gonzaga

du

São Carlos – São Paulo
2000

USP/IFSC/SBI

8-2-001363

IFSC-USP SERVIÇO DE BIBLIOTECA
INFORMAÇÃO

Neves, Evelina Maria de Almeida

"Estratégia atencional para busca visual e reconhecimento invariante de objetos baseada na integração de características bottom-up e top-down " / Evelina Maria de Almeida Neves – São Carlos, 2000

Tese (Doutorado) – Instituto de Física de São Carlos da Universidade de São Paulo, 2000 – páginas: 283

Área: Física Aplicada-opção: Física Computacional

Prof. Dr. Adilson Gonzaga

1. Atenção visual, 2. Busca visual, 3. Lógica Fuzzy, 4. Gestalt

I. Título

UNIVERSIDADE
DE SÃO PAULO

Instituto de Física de São Carlos

Av. Dr. Carlos Botelho, 1465
CEP 13560-250 - São Carlos - SP
Brasil

Fone (016) 273-9333

Fax (016) 272-2218

MEMBROS DA COMISSÃO JULGADORA DA TESE DE DOUTORADO DE EVELINA MARIA DE ALMEIDA NEVES APRESENTADA AO INSTITUTO DE FÍSICA DE SÃO CARLOS, DA UNIVERSIDADE DE SÃO PAULO, EM 30 DE JUNHO DE 2000.

COMISSÃO JULGADORA:

Prof. Dr. Adilson Gonzaga/EESC-SEL-USP

Profa. Dra. Maria Stela Veludo de Paiva/EESC-SEL-USP

Prof. Dr. Luciano da Fontoura Costa/IFSC-FFI-USP

Prof. Dr. Roberto Marcondes Cesar Júnior/IME-USP

Prof. Dr. Clésio Luis Tozzi/UNICAMP

e-mail: wladerez@ifsc.usp.br

Ao Márcio e Yasmin, com Amor

“Só se vê bem com o coração, o essencial é invisível aos olhos....”
Anthony Saint Exupery

AGRADECIMENTOS

- Ao Prof. Dr. Adilson Gonzaga, pela amizade, orientação, oportunidade e pelos valiosos ensinamentos que ultrapassaram as fronteiras deste trabalho.
- Ao João Borelli, sou eternamente grata pela amizade, apoio nas horas difíceis e inestimável colaboração.
- Ao Profs. Evandro L.L. Rodrigues e Maria Stela Veludo Paiva, pela amizade demonstrada.
- À Prof. Dr. Annie F.F. Slaets por seu inestimável auxílio.
- À toda a minha família pelo apoio e incentivo para realização de mais este trabalho.
- À todos os meus amigos, em especial à Ana Cláudia, Anderson, Annete, Carl, Celso, Fátima, Leo, Luiz, Maurício, Mônaco, Patrícia, Renata, Rodrigo, Xandão e Zé Luiz, pela amizade, companheirismo e pelos momentos agradáveis que passamos juntos.
- À solicitude incondicional das bibliotecárias do IFSC – USP, em especial à Mara e à Neusa, cujos préstimos a nós dedicados transcendem as suas obrigações profissionais.
- À todos os funcionários do IFSC e da EESC, e em especial à Wladerez pela consideração em todos os momentos.
- Ao Durval, Beatriz, Eduardo e Maurício, pelos anos maravilhosos que tenho vivido na Fazenda Paraizo e que me proporcionaram a tranquilidade para concluir este trabalho.
- Ao CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico) pela concessão da bolsa de estudos.
- À Deus, por estar sempre presente ao meu lado, possibilitando a conclusão de mais uma etapa da minha vida.

SUMÁRIO

LISTA DE FIGURAS	i
LISTA DE TABELAS	xiii
RESUMO	xvii
ABSTRACT	xix
CAPÍTULO 1 – VISÃO GERAL	1
1 – INTRODUÇÃO	1
2 - MOTIVAÇÕES TEÓRICAS	3
3 – OBJETIVOS	5
4 – DESCRIÇÃO GERAL DO TRABALHO	6
5 - ORGANIZAÇÃO DA DISSERTAÇÃO	12
CAPÍTULO 2 – PSICOLOGIA DA FORMA	15
1 – INTRODUÇÃO	15
2 – TEORIAS DE PERCEPÇÃO HUMANA DE FORMA	17
2.1 – Abordagem Estruturalista	18
2.2 – Abordagem Gestalt – 1910	19
2.2.1 - Leis de organização	21
2.2.2 - Dependência funcional da figura e fundo	28
2.3 – Teoria do comportamento de Hebb (1949)	29
2.4 – Abordagem Ecológica de Gibson (1966)	29
2.5 – Processamento Perceptivo	31
2.6 – Abordagem Construtivista (1970)	33
2.7 – Teoria Computacional de Marr (1976)	33
2.8 – Teoria Psicológica de Biederman (1987)	34
2.9 - Teoria da percepção da Forma de Treisman (1986)	36
2.10 – Rotinas Visuais (1984)	41
2.10.1 – Rotinas visuais e seus papéis no processamento da informação visual	46
2.10.2 – Operações elementares	50
2.11 – Teorias baseadas em Esquemas	55
2.12 – Percepção canônica dos objetos	56
3 - EXTENDENDO A PSICOLOGIA GESTALT	57
3.1 – Gestalt e Atenção	60
3.2–Organização da Percepção em Visão Computacional	63

4 – CONCLUSÕES	68
CAPÍTULO 3 - ATENÇÃO VISUAL HUMANA	72
1 – INTRODUÇÃO	72
2 – MOTIVOS PARA A EXISTÊNCIA DA ATENÇÃO VISUAL	74
3 – FUNÇÕES DO SISTEMA ATENCIONAL	78
4 - DEFINIÇÃO DE CONCEITO LIMITE	80
5 - TIPOS DE ATENÇÃO	81
6 - RELAÇÃO DA ATENÇÃO COM O MOVIMENTO DOS OLHOS	84
7 - BUSCA VISUAL	87
8 - MEMÓRIA E ATENÇÃO	92
9 - DESENVOLVIMENTO DA ATENÇÃO VISUAL	96
9.1 – Exploração e pesquisa	98
9.2 - Predisposição e Expectativa	99
9.3 - Divisão e flutuação da atenção	101
9.4 – Efeitos na variação e mudança na Percepção	103
10 – MODELOS TEÓRICOS PARA A ATENÇÃO VISUAL HUMANA	103
10.1 - O modelo da Lente de Zoom para a Atenção (<i>Zoom –Lens Model</i>)	105
10.2 - O modelo do foco de atenção (<i>Spotlight Model</i>)	105
10.3 – A teoria dos <i>textons</i> de Julesz	106
10.4 - Teoria da Integração de características de Treisman (<i>Feature Integration</i>)	106
CAPÍTULO 4 - ATENÇÃO VISUAL COMPUTACIONAL	107
1 – INTRODUÇÃO	107
2 – MODELOS COMPUTACIONAIS	107
2.1 – O Mapa de Saliência (<i>Saliency Map</i>)	108
2.2 – O modelo de Chapman	109
2.3 - O modelo da Busca Guiada (<i>Guided Search</i>)	110
2.4 - O VISIT de Ahmad	112
2.5 - O Neocognitron	113
2.6 - O modelo de Mozer	115
2.7 – O modelo de Sandon	115
2.8 – O modelo de Milanese	115
3 – CONCLUSÕES	116

CAPÍTULO 5 - ESTRATÉGIA ATENCIONAL PARA BUSCA VISUAL E RECONHECIMENTO INVARIANTE DE OBJETOS ATRAVÉS DA INTEGRAÇÃO DE CARACTERÍSTICAS <i>BOTTOM-UP</i> E <i>TOP-DOWN</i>	121
1 - VISÃO GERAL DO MODELO PROPOSTO	121
2 – APRESENTAÇÃO E DISCUSSÃO DOS FUNDAMENTOS DO MODELO	123
2.1 – Aquisição das imagens	124
2.2 – Segmentação da Imagem	126
2.3 – Escolha dos atributos	128
2.4 – Representação das características em múltiplos mapas	135
2.5 – Mapas de Características Primários x Intermediários	140
2.6 – Mapas de Conspicuidade	143
2.6.1 - Mapa de conspicuidade de distância	145
2.6.2 – Mapa de conspicuidade de média de nível de cinza	152
2.6.3 – Mapa de conspicuidade de Formato	153
2.6.4 – Mapa de Conspicuidade de orientação	155
2.6.5 – Mapa de Conspicuidade de tamanho	156
2.7 – Conceitos Fuzzy dentro do modelo proposto	157
2.7.1 – Os conjuntos Fuzzy	159
2.7.2 – Extração do Conhecimento	161
2.7.2.1 – Determinação das funções de pertinência do subsistema <i>Bottom-up</i>	162
2.7.2.1.1 – Determinação das Funções de pertinência para o conjunto <i>fuzzy</i> conspicuidade de tamanho	162
2.7.2.1.2 – Determinação das Funções de pertinência para o conjunto <i>fuzzy</i> “Conspicuidade de cinza”	168
2.7.2.1.3 – Determinação das Funções de pertinência para o conjunto <i>fuzzy</i> “Conspicuidade de orientação”	172
2.7.2.1.4 – Determinação das Funções de pertinência para o conjunto <i>fuzzy</i> “Conspicuidade de distância”	176
2.7.2.1.5 – Determinação das Funções de pertinência para o conjunto <i>fuzzy</i> “Conspicuidade de Formato”	180
2.7.2.2 – Determinação das Regras de Inferência <i>fuzzy</i> para os conjuntos <i>fuzzy</i> de “Conspicuidade ”	181
2.7.2.3 – Inclusão de informações <i>Top-down</i>	193

2.7.2.4 – Determinação das funções de pertinência do subsistema <i>top-down</i>	195
2.7.2.4.1 - Determinação das funções de pertinência para o conjunto <i>Fuzzy</i> “ <i>nível de cinza</i> ” do subsistema <i>top-down</i>	195
2.7.2.4.2 – Determinação das funções de pertinência para o conjunto <i>Fuzzy</i> “ <i>Tamanho</i> ” do subsistema <i>top-down</i>	199
2.7.2.4.3 – Determinação das funções de pertinência para o conjunto <i>Fuzzy</i> “ <i>Orientação</i> ” do subsistema <i>top-down</i>	204
2.7.2.4.4 – Determinação das funções de pertinência para o conjunto <i>Fuzzy</i> “ <i>Distância</i> ” do subsistema <i>top-down</i>	207
2.7.2.4.5 – Determinação das funções de pertinência para o conjunto <i>fuzzy</i> “ <i>Forma</i> ” do subsistema <i>top-down</i>	212
2.7.2.5 – Determinação das Regras de Inferência <i>fuzzy</i> para o subsistema <i>top-down</i>	212
2.7.2.6– O subsistema <i>Fuzzy</i> de Integração	220
2.8 – Discussão dos Resultados	236
 CAPÍTULO 6 – CONCLUSÕES	238
1 – CONSIDERAÇÕES FINAIS	238
2 – EXTENSÃO DO MODELO BÁSICO E SUGESTÕES PARA TRABALHOS FUTUROS	240
3 – CONTRIBUIÇÕES E CONCLUSÕES FINAIS	242
 REFERÊNCIAS BIBLIOGRÁFICAS	251
 APÊNDICE 1 – FISILOGIA DA FORMA	284
 APÊNDICE 2 – FORMA COMPUTACIONAL BASEADA EM MOMENTOS	316
 APÊNDICE 3 – LÓGICA <i>FUZZY</i>	367
 APÊNDICE 4 – TABELAS DE REGRAS DEFINIDAS PARA AS TAREFAS DO SUBSISTEMA TOP-DOWN	399

LISTA DE FIGURAS

Figura 1.1 - Áreas de atuação de Reconhecimento de Padrões e sua relação com as diversas subáreas do conhecimento.	2
Figura 1.2 – Diagrama em blocos do Modelo Proposto	11
Figura 2.1 – Diversas representações para o objeto “cadeira”	16
Figura 2.2 - Exemplos de figuras vistas preferencialmente como globais	20
Figura 2.3 - Formas que podem ser vistas de diversas maneiras.	22
Figura 2.4 - Formas pretas e brancas que se tornam perceptivamente organizadas em um dalmata	23
Figura 2.5 - Ilustração da Lei da vizinhança	23
Figura 2.6 - Ilustração da Lei da igualdade	24
Figura 2.7 - Ilustração da Lei da forma fechada	24
Figura 2.8 - Ilustração da Lei do destino comum	25
Figura 2.9 - Exemplo da lei da boa figura.	25
Figura 2.10 – Exemplo da lei da Simetria	26
Figura 2.11 – Figura –fundo	29
Figura 2.12 – Exemplo de como um processo cognitivo pode ser quebrado em uma sequência de passos.	32
Figura 2.13 – Exemplos de <i>geons</i> (categorias de formas 3-D) proposta por Biederman, (1987) para descrever partes dos objetos	35

Figura 2.14 – É mais fácil determinar “Q” em (a) mas é difícil determinar “T” em (b)	36
Figura 2.15 - Experiências realizadas por Treisman para explicar a busca de uma alvo	37
Figura 2.16 – Busca Visual do estímulo R não é instantânea, a menos que ele seja encontrado primeiro ao acaso.	38
Figura 2.17 – Distinção imediata (<i>pop-out boundary</i>) entre regiões de “Os” e “Vs”.	38
Figura 2.18 – (a) A orientação é uma força poderosa responsável pelo agrupamento perceptivo	39
Figura 2.19 – Diagrama simplificado da sequência de processamento para a teoria de Treisman	40
Figura 2.20 – Modelo hipotético proposto por Anne Treisman (1986)	41
Figura 2.21 – Desenhos esquemáticos de uma face (a) normalmente disposta e (b) “embaralhada”.	42
Figura 2.22 – A figura global pode ser traçada sem reconhecer todos os seus componentes individuais	55
Figura 2.23 – Contorno subjetivo.	65
Figura 3.1 - Leva-se menos tempo para detectar uma linha oblíqua entre linhas verticais (a) do que detectar uma linha vertical entre linhas oblíquas (b).	90
Figura 3.2 - Efeito do agrupamento em tarefas de busca visual	91
Figura 3.3 - Forma ambígua	101
Figura 3.4 - Alternância de figura e fundo	102
Figura 3.5 – Cruz de Malta	102

Figura 3.6 - “Minha esposa e minha sogra”	104
Figura 4.1 – Pirâmide atencional de Koch e Ullman.	109
Figura 4.2 – Modelo de Busca Guiada de Cave & Wolfe	111
Figura 4.3 – Visão Geral do VISIT para busca visual	113
Figura 5.1 – Sistema de aquisição de imagem	124
Figura 5.2 – Cena com objetos	125
Figura 5.3 – Distribuição dos valores obtidos para o invariante de Hu “ ϕ_1 ” para 12 cenas	130
Figura 5.4 – Distribuição dos valores obtidos para o invariante de Hu “ ϕ_2 ” para 12 cenas	130
Figura 5.5 – Distribuição dos valores obtidos para o invariante de Hu “ ϕ_3 ” para 12 cenas	131
Figura 5.6 – Distribuição dos valores obtidos para o invariante de Hu “ ϕ_4 ” para 12 cenas	131
Figura 5.7 – Distribuição dos valores obtidos para o invariante de Hu “ ϕ_5 ” para 12 cenas	131
Figura 5.8 – Distribuição dos valores obtidos para o invariante de Hu “ ϕ_6 ” para 12 cenas	132
Figura 5.9 – Distribuição dos valores obtidos para o invariante de Hu “ ϕ_7 ” para 12 cenas	132
Figura 5.10 - Distribuição dos valores obtidos para a soma dos invariantes de Hu “ $\phi_1+\dots+\phi_7$ ” para 12 cenas	133

Figura 5.11 – Mapa de saliência global	134
Figura 5.12 – Mapa de saliência global (GSM) baseado no modelo de Cave & Wolfe (1990)	138
Figura 5.13 – Exemplo de uma imagem (A) e as tarefas (B) determinar o menor círculo, (C) encontrar os círculos, (D) encontrar os menores objetos	139
Figura 5.14 – Exemplos de mapas de característica de nível intermediário	142
Figura 5.15 – Extração de mapas de características e dos mapas de conspicuidade e suas relações com a base de dados construído a partir dos conceitos da lógica <i>fuzzy</i>	145
Figura 5.16 – Cena contendo um objeto mais distante em relação ao grupo	146
Figura 5.17 – Cena contendo objetos em diferentes posições	149
Figura 5.18- Cena contendo objetos aproximadamente eqüidistantes	150
Figura 5.19 – Cena contendo um objeto conspícuo com relação à média de nível de cinza	152
Figura 5.20 – Cena contendo objetos com diferentes formatos	154
Figura 5.21 – Cena com objetos contendo orientações conspícuas	155
Figura 5.22 – Cena contendo objetos de tamanhos diferentes	156
Figura 5.23 – Visão Geral do modelo proposto	159
Figura 5.24 – Figura análoga à utilizada no teste psicofísico para medida do grau de atenção em função da diferença de tamanho dos objetos, com relação ao primeiro objeto	163
Figura 5.25 – Gráfico resultante da avaliação do questionário para o termo “pouca conspicuidade” de tamanho	164

- Figura 5.26 – Gráfico resultante da avaliação do questionário para o termo “Média conspicuidade” de tamanho 165
- Figura 5.27 – Gráfico resultante da avaliação do questionário para o termo “Muita conspicuidade” de tamanho 166
- Figura 5.28 – Funções de Pertinência para os termos “Pouca”, “Média” e “Muita” do conjunto *fuzzy* de entrada “Conspicuidade de Tamanho” em função dos tamanhos normalizados 166
- Figura 5.29 – Funções de Pertinência para os termos “Pouca”, “Média” e “Muita” do conjunto *fuzzy* de entrada “Conspicuidade de Tamanho” em função da conspicuidade normalizada no intervalo [0 1] 167
- Figura 5.30 – Aproximação das funções de pertinência do conjunto *fuzzy* “conspicuidade de tamanho” 167
- Figura 5.31 - Figura análoga à utilizada no teste psicofísico para medida do grau de atenção em função da diferença de tonalidade dos objetos em relação ao primeiro objeto 168
- Figura 5.32 - Gráfico resultante da avaliação do questionário para o termo “Pouca conspicuidade” de nível de cinza. 169
- Figura 5.33 - Gráfico resultante da avaliação do questionário para o termo “Média conspicuidade” de nível de cinza. 169
- Figura 5.34 - Gráfico resultante da avaliação do questionário para o termo “Muita conspicuidade” de nível de cinza. 170
- Figura 5.35 – Representação das funções de pertinência para os termos “Pouca”, “Média” e “Muita” conspicuidade do conjunto *fuzzy* “Conspicuidade de Nível de Cinza” em função dos níveis de cinza normalizados 171

- Figura 5.36 - Representação das funções de pertinência para os termos “Pouca”, “Média” e “Muita” conspicuidade do conjunto *fuzzy* “Conspicuidade de Nível de Cinza” em função da conspicuidade normalizada entre [0 1] 171
- Figura 5.37 – Aproximação das funções de pertinência dos termos “Pouca”, “Média” e “Muita” do conjunto *fuzzy* de entrada “conspicuidade de Nível de Cinza” 172
- Figura 5.38 - Figura análoga à utilizada no teste psicofísico para medida do grau de conspicuidade em função da inclinação dos objetos, em relação ao primeiro objeto 172
- Figura 5.39 – Gráfico resultante da pesquisa realizada para o termo “pouca” conspicuidade de orientação 173
- Figura 5.39 – Gráfico resultante da pesquisa realizada para o termo “pouca” conspicuidade de orientação 173
- Figura 5.40 – Gráfico resultante da pesquisa realizada para o termo “média” conspicuidade de orientação 174
- Figura 5.41 – Gráfico resultante da pesquisa realizada para o termo “muita” conspicuidade de orientação 174
- Figura 5.42 - Representação das funções de pertinência para os termos “Pouca”, “Média” e “Muita” conspicuidade do conjunto *fuzzy* “Conspicuidade de orientação” em função das orientações normalizadas 175
- Figura 5.43 - Representação das funções de pertinência para os termos “Pouca”, “Média” e “Muita” conspicuidade do conjunto *fuzzy* “Conspicuidade de Orientação” em função da conspicuidade normalizada entre [0 1] 175
- Figura 5.44 - Aproximação das funções de pertinência para os termos “Pouca”, “Média” e “Muita” do conjunto *fuzzy* de entrada “conspicuidade de orientação” 175
- Figura 5.45 - Figura análoga à utilizada no teste psicofísico para medida do grau de conspicuidade em função da distância dos objetos em relação ao primeiro objeto 176

- Figura 5.46 – Gráfico resultante da pesquisa realizada para o termo “pouca” conspicuidade de distância 177
- Figura 5.47 – Gráfico resultante da pesquisa realizada para o termo “Média” conspicuidade de distância. 178
- Figura 5.48 – Gráfico resultante da pesquisa realizada para o termo “Muita” conspicuidade de distância 178
- Figura 5.49 - Representação das funções de pertinência para os termos “Pouca”, “Média” e “Muita” conspicuidade do conjunto *fuzzy* “Conspicuidade de tamanho” em função das distâncias normalizadas 179
- Figura 5.50 - Representação das funções de pertinência para os termos “Pouca”, “Média” e “Muita” conspicuidade do conjunto *fuzzy* “Conspicuidade de Distância” em função das conspicuidades normalizadas no intervalo [0 1] 179
- Figura 5.51 - Aproximação das funções de pertinência dos termos “Pouca”, “Média” e “Muita” do conjunto *fuzzy* “Conspicuidade de Distância” 180
- Figura 5.52 – Função de Pertinência adotadas para o conjunto *fuzzy* “conspicuidade de formato” 181
- Figura 5.53 – Exemplo mostrando como a configuração global de uma cena pode ser dependente do contexto das características presentes na mesma 182
- Figura 5.54 – Regras de inferência especificadas por um domínio especialista e por literatura relacionada 183
- Figura 5.55 – funções de pertinência para as variáveis “pouca” (p), “média-pouca” (mp), “média” (m), “média-muita” (mg) e “muita” (g) do conjunto *fuzzy* de saída “saliência *bottom-up*” 184
- Figura 5.56 – Exemplo de cena com objetos que apresentam diferentes características 189
- Figura 5.57 – Exemplo de cena onde os objetos compartilham diferentes características 190

Figura 5.58 – Exemplo de cena contendo objetos com diferentes características	190
Figura 5.59 – Cena contendo objetos com pouca conspicuidade entre si	191
Figura 5.60 – Exemplo de uma cena contendo um objeto que difere dos outros objetos em todas as dimensões de características	191
Figura 5.61 – Cena contendo um objeto com maior conspicuidade de média de nível de cinza e outro objeto com conspicuidade de orientação em relação aos demais objetos presentes	192
Figura 5.62 – Cena contendo alfinetes e pregos.	192
Figura 5.63 – Faixa de nível de cinza compreendendo os valores de 0 a 255	195
Figura 5.64 – Gráfico resultante da avaliação do questionário para o termo “pouco nível de cinza” (“escuro”)	196
Figura 5.65 – Gráfico resultante da avaliação do questionário para o termo “Médio Nível de Cinza”(tonalidade média)	197
Figura 5.66 – Gráfico resultante da avaliação do questionário para o termo “ <i>Large Gray</i> ”(claro)	198
Figura 5.67 – Representação gráfica das funções de pertinência (<i>membership functions</i>) do conjunto <i>fuzzy</i> “nível de cinza” (“gray”) para os termos “pouco” (“small”), “médio” (“medium”) e muito (“large”)	198
Figura 5.68 – Funções de pertinência para o conjunto <i>fuzzy</i> “nível de cinza” (“Gray”)	199
Figura 5.69 - Figura compreendendo objetos com tamanhos normalizados, utilizada como referência para a determinação das funções de pertinências para os atributos “pequeno”, “médio” e “grande” do conjunto <i>fuzzy</i> “tamanho”	200

Figura 5.70 - Gráfico resultante da avaliação do questionário para o termo “pequeno” do conjunto <i>fuzzy</i> tamanho.	201
Figura 5.71 - Gráfico resultante da avaliação do questionário para o termo “tamanho médio”	201
Figura 5.72 - Gráfico resultante da avaliação do questionário para o termo “tamanho grande”	202
Figura 5.73 – Representação gráfica das funções de pertinência do conjunto <i>fuzzy</i> “Tamanho” para os termos “pequeno” (<i>small</i>), “médio” (<i>medium</i>) e muito (“ <i>large</i> ”)	203
Figura 5.74 – Funções de pertinência para o conjunto <i>fuzzy</i> “Tamanho”	203
Figura 5.75 - Figura compreendendo objetos com tamanhos normalizados, utilizada como referência para a determinação das funções de pertinências para os atributos “pouco” e “muito” do conjunto <i>fuzzy</i> “orientação”	204
Figura 5.76 - Gráfico resultante da avaliação do questionário para o termo “pouco inclinado”	205
Figura 5.77 - Gráfico resultante da avaliação do questionário para o termo “muito inclinado”	206
Figura 5.78 - Representação gráfica das funções de pertinência (<i>membership functions</i>) do conjunto <i>fuzzy</i> “Orientação” para os atributos “pouco” (<i>small</i>) e muito (“ <i>large</i> ”) inclinado	206
Figura 5.79 – Funções de pertinência aproximadas para o conjunto <i>fuzzy</i> “Orientação”	207
Figura 5.80 - Figura compreendendo objetos com tamanhos normalizados, utilizada como referência para a determinação das funções de pertinências para os atributos “pouca”, “média” e “muita” do conjunto <i>fuzzy</i> “Distância”	208

Figura 5.81 - Gráfico resultante da avaliação do questionário para o termo “Pouca Distância” (perto)	209
Figura 5.82 - Gráfico resultante da avaliação do questionário para “Média Distância”	210
Figura 5.83 – Gráfico resultante da avaliação do questionário para “Longe” (“muita distância”)	210
Figura 5.84 - Representação gráfica das funções de pertinência (<i>membership functions</i>) para os termos “pouca” (<i>small</i>), “média” (“ <i>medium</i> ”) e muita (“ <i>large</i> ”) distância	211
Figura 5.85 – Funções de pertinência para o conjunto <i>fuzzy</i> “ <i>Distância</i> ”	211
Figura 5.86 - Inferência chave para as tarefas envolvendo formas no sub-sistema <i>top-down</i> .	212
Figura 5.87 – Funções de pertinência para as variáveis “pouca” (p), “média-pouca” (mp), “média” (m), “média-muita” (mg) e “muita” (g) do conjunto <i>fuzzy</i> de saída “saliência <i>top-down</i> ”	213
Figura 5.88 – Cena contendo objetos com várias tonalidades	214
Figura 5.89 – Cena contendo objetos de diferentes tamanhos	215
Figura 5.90 – Cena contendo objetos com diferentes formatos	216
Figura 5.91 – Cena contendo diferentes objetos com diferentes tamanhos e tonalidades	217
Figura 5.92 – Cena contendo objetos (alicates e grifos) com diferentes tonalidades e inclinações	218
Figura 5.94 – Cena contendo objetos com diferentes características	218
Figura 5.95 – Funções de pertinência para as variáveis “pouca” (p), “média-pouca” (mp), “média” (m), “média-muita” (mg) e “muita” (g) do conjunto <i>fuzzy</i> de saída “saliência <i>top-down</i> ”	221

Figura 5.96 – Cena com objetos utilizada na tarefa de “determinar o círculo grande” e na tarefa “determinar o círculo pequeno”	222
Figura 5.97 – Mapas de Conspicuidades de Média de Nível de Cinza, tamanho, orientação, formato e distância calculados para os objetos da figura 5.96	223
Figura 5.98 – Índices medidos para a cena mostrada na figura 5.96, para a tarefa de determinar os “círculos grandes”	223
Figura 5.99 – Índices medidos para a cena mostrada na figura 5.96, para a tarefa de determinar os “círculos pequenos”	224
Figura 5.100 – Cena análoga à mostrada na figura 5.96, onde o objeto alvo (círculo pequeno) apresenta uma contribuição <i>bottom-up</i> maior devido à diferença de formato com os demais objetos presentes na cena	225
Figura 5.101 – Mapas de Conspicuidade calculados para os objetos presentes na figura 100.	225
Figura 5.102 – Índices medidos para a cena mostrada na figura 5.100, para a tarefa de determinar os “círculos pequenos”	226
Figura 5.103 – Exemplo de uma cena para demonstrar como a presença de um objeto na cena pode interferir na execução de uma tarefa	227
Figura 5.104 – Mapas de Conspicuidades de Média de Nível de Cinza, tamanho, orientação, formato e distância calculados para os objetos da figura 5.103.	227
Figura 5.105 – Comparação dos índices de saliência obtidos nos subsistemas <i>bottom-up</i> , <i>top-down</i> e de integração para a cena mostrada na figura 5.103	228
Figura 5.106 – Cena contendo chave inglesa e alicate	229
Figura 5.107 – Mapa de conspicuidade obtido para os objetos da cena mostrada na figura 5.104	229

Figura 5.108 – Cena análoga à mostrada na figura 5.106 onde o alvo (grifo) está presente.	231
Figura 5.109 – Mapas de Conspicuidade para a cena 5.108	231
Figura 5.110 – Cena contendo alicates que compartilham diferentes características	232
Figura 5.111 – Mapa de conspicuidade para a cena mostrada na figura 5.110.	233
Figura 5.112 – Índices de saliência <i>bottom-up</i> , <i>top-down</i> e <i>global</i> , obtidos para a tarefa de determinar os alicates escuros na cena mostrada na figura 5.110	234
Figura 5.113 – Imagem contendo chaves inglesas com diferentes características	234
Figura 5.114 – Mapas de Conspicuidade medido para os objetos presentes na cena mostrada na figura 5.113	235

LISTA DE TABELAS

Tabela 5.1 – Matriz de saliência de distância para a cena mostrada na figura 5.16	147
Tabela 5.2 – Matriz de saliência para objetos da figura 5.16, normalizada em relação a $T=0.3d_{max}$	148
Tabela 5.3 – matriz de saliência de distância para a cena mostrada na figura 5.17	149
Tabela 5.4 – Matriz de saliência para figura 5.17 normalizada com relação a $T=0.3\%$	150
Tabela 5.5 – matriz de saliência de distância para a cena mostrada na figura 5.18.	151
Tabela 5.6 – Matriz de saliência para objetos da figura 5.18, normalizada em relação a $T=0.3d_{max}$	151
Tabela 5.7 – matriz de saliência de média de nível de cinza para a cena mostrada na figura 5.19	152
Tabela 5.8 – Matriz de saliência de média de nível de cinza normalizada para a cena mostrada na figura 5.19.	153
Tabela 5.9 – Matriz de saliência de formato para a cena mostrada na figura 5.20	154
Tabela 5.10 – Matriz de saliência de formato normalizada para a cena mostrada na figura 5.20.	155
Tabela 5.11 – Matriz de saliência de orientação para a cena mostrada na figura 5.21	156
Tabela 5.12 – Matriz de saliência de tamanho para a cena mostrada na figura 5.22	157
Tabela 5.13 – Conjuntos <i>Fuzzy</i> da primeira camada da rede	160
Tabela 5.14 – Conjuntos <i>fuzzy</i> para a segunda camada da rede	160

Tabela 5.15 – Respostas obtidas para a característica “Pouca conspicuidade” de tamanho	163
Tabela 5.16 – Respostas obtidas para a característica “Média conspicuidade” de tamanho	164
Tabela 5.17 – Respostas obtidas para a característica “Muita conspicuidade” de tamanho	165
Tabela 5.18 – Respostas obtidas para a característica “pouca conspicuidade” de Nível de Cinza	168
Tabela 5.19 – Respostas obtidas para a característica “média conspicuidade” de Nível de Cinza	169
Tabela 5.20 – Respostas obtidas para a característica “Muita conspicuidade” de Nível de Cinza	170
Tabela 5.21 – Respostas obtidas para “pouca”, “média” e “muita” conspicuidade de orientação	173
Tabela 5.22 – Respostas obtidas para “pouca”, “média” e “muita” conspicuidade de distância	177
Tabela 5.23 – Tabela de regras de inferência para determinação do índice de saliência <i>bottom-up</i>	185
Tabela 5.24 – Respostas obtidas para a característica “escuro”	196
Tabela 5.25 – Respostas obtidas para a característica “tonalidade média”	197
Tabela 5.26 – Respostas obtidas para a característica “claro”	197
Tabela 5.27 – Respostas obtidas para a característica “Pequeno”	200
Tabela 5.28 – Respostas obtidas para a característica “tamanho médio”	201

Tabela 5.29 – Respostas obtidas para a característica “Tamanho Grande”	202
Tabela 5.30 – Respostas obtidas para a característica “pouco inclinado”	205
Tabela 5.31 – Respostas obtidas para a característica “muito inclinado”	205
Tabela 5.32 – Respostas obtidas para a característica “Perto”	208
Tabela 5.33 – Respostas obtidas para a característica “média distância”	209
Tabela 5.34 – Respostas obtidas para a característica “longe”	210
Tabela 5.35 – Regras de Inferência para determinação do índice de Saliência Global”	220
Tabela 5.36 – Comparação entre os índices de saliência <i>bottom-up</i> , <i>top-down</i> e <i>global</i> , obtidos para os objetos da figura 5.96 para a tarefa de determinar o “círculo pequeno”	223
Tabela 5.37 – Comparação entre os índices de saliência <i>bottom-up</i> , <i>top-down</i> e <i>global</i> obtidos para os objetos da figura 5.97 para a tarefa de determinar o “círculo grande”	223
Tabela 5.38 – Comparação entre os índices de saliência <i>bottom-up</i> , <i>top-down</i> e <i>global</i> , obtidos para os objetos da figura 5.100 para a tarefa de determinar o “círculo pequeno”	226
Tabela 5.39 – Comparação entre os índices de saliência para os objetos da figura 5.103	227
Tabela 5.40 – Comparação entre os índices de saliência obtidos para os objetos da figura 5.103	228
Tabela 5.41 – Índices obtidos para os objetos da figura 5.106, em diversas tarefas.	230
Tabela 5.42 – Índices obtidos para a cena mostrada na figura 5.106, para a tarefa de determinar o grifo, escuro e inclinado	232
Tabela 5.43 – Índices obtidos para a cena mostrada na figura 5.110, para a tarefa de determinar o alicate escuro	233

Tabela.5.44 – Índices obtidos para a cena mostrada na figura 113, para a tarefa de determinar chave inglesa clara, pequena e mais distante

RESUMO

Uma das tarefas básicas dos mecanismos atencionais é decidir qual localização dentro do campo visual devemos prestar atenção primeiro. Um objeto que contenha características distintas (tais como orientação, forma, cor, tamanho, brilho, textura, etc.. diferentes) pode atrair a atenção de uma maneira *bottom-up*. A informação *top-down* baseia-se no conhecimento prévio e tem uma grande influência nas localizações atendidas. Inspirado nos mecanismos da Atenção Visual Humana, embora sem a pretensão de simulá-la, este trabalho prevê o desenvolvimento de uma nova metodologia que integra os dois tipos de informações: *bottom-up* e *top-down*. Características *bottom-up* são geradas a partir de Momentos e essas informações são utilizadas em mapas de saliência, enquanto que um conhecimento prévio é utilizado para gerar pistas *top-down*.

Neste trabalho, desenvolveu-se uma metodologia específica para a busca e o reconhecimento visual em cenas com múltiplos objetos, utilizando para isso uma rede *fuzzy* contendo três subsistemas *fuzzy*. Dada uma imagem de entrada, o objetivo consiste em se detectar regiões que possam conter informações mais significativas, a fim de que se possa guiar e restringir processamentos mais complexos. A inclusão de mecanismos de atenção (seleção de uma região de interesse dentro da imagem) é de fundamental importância pois os resultados obtidos pelo modelo podem ser usados para controlar a aquisição da imagem de uma maneira dinâmica.

O modelo proposto está estruturado em três estágios principais: O primeiro estágio consiste em se segmentar os objetos e extrair características globais dos mesmos baseadas principalmente na teoria dos momentos, tais como tamanho, orientação, formato e distância e também média de nível de cinza. Através da comparação de um objeto com os outros presentes na cena, características *bottom-up* de conspicuidade são usadas para guiar a atenção ao objeto mais diferente. Através do uso da lógica *fuzzy* é possível inferir com grande flexibilidade algumas regras de decisão baseadas nos princípios de percepção visual tais como as leis Gestalt. O segundo estágio consiste de um subsistema *fuzzy top-down* que combina diferentes características de acordo com a relevância das mesmas em diferentes tarefas. Finalmente o terceiro estágio consiste de um subsistema *fuzzy* que integra as informações obtidas dos subsistemas anteriores e fornece um índice geral de saliência, e indica a provável localização do objeto a ser reconhecido. A nova abordagem foi testada

com objetos geométricos levando-se em consideração as características que atraem a atenção dos seres humanos.

Palavras Chaves: Atenção Visual, Busca Visual, Lógica *Fuzzy*, Teoria dos Momentos, Gestalt, Percepção Visual, Seleção, Processamento *Bottom-up* e *Top-down*, Mapas de Saliência, Reconhecimento Invariante, Integração de Características.

ABSTRACT

One of the basic tasks assigned to the human attentional mechanisms is to decide which location in the visual field we must pay attention first. An object containing distinctive features (such as different orientation, shape, color, size, shine, texture, etc.) can attract attention in a bottom-up way. Top-down information is based on the previous knowledge and has a large influence on the attended locations. Inspired on human visual attention mechanisms, although it doesn't want simulate it, this work presents a new methodology to integrate two different kind of information: bottom-up and top-down. Bottom-up features are obtained from Moment Theory and these information are used in salience maps, while a previous knowledge is used to create top-down hints.

In this work, an specific methodology to visual search and recognition was developed to be applied to scenes containing multiple objects by a fuzzy net with three fuzzy subsystems. The aim of this methodology is to detect regions that may contain the most significant information, in order to guide and to restrict most complex processing. The inclusion of attentional mechanisms (the selection of a region of interest in the image) is fundamental and can be used to control the image acquisition in a dynamic way.

The proposed model is structured in three main stages. The first stage segments the objects and extracts global features of them, based on the Moment Theory such as size, orientation, shape and distance and gray level average. By comparing one object with the other ones present in the scene, bottom-up features of conspicuity are used to guide the attention to the most different object. The Fuzzy Logic allow us to infer with great flexibility some of decision rules based on the visual perception principles such as the Gestalt Laws. The second stage is a top-down fuzzy subsystem that combines different features according to the relevance of them in different tasks. Finally, the third stage is a fuzzy subsystem that integrates the information obtained from the previous subsystems and gives us a general salience index. The new methodology was tested in geometrical objects considering the features that attracts attention to human beings.

Key Words: Visual Attention, Visual Search, Perception, Fuzzy Logic, Moment Theory, Selection, Bottom-up Processing, Top-down processing, Salience Maps, Invariant Recognition, Feature Integration.

CAPÍTULO 1

VISÃO GERAL

1 – INTRODUÇÃO

Desde os primeiros interesses científicos do homem pelo funcionamento do próprio corpo a percepção do mundo exterior desperta especial atenção. Através de um processo dialético de observação a técnica evoluiu para máquinas que tentam imitar não apenas os mecanismos biológicos mas também seu comportamento psicológico.

Na busca do homem pela criação, diversas disciplinas têm surgido tais como Inteligência Artificial, Processamento de Imagens, Robótica e outras que abrangem grandes áreas e subáreas do conhecimento humano. O nome da disciplina que abriga a teoria da percepção aplicada às diversas áreas da Ciência é o Reconhecimento de Padrões. Essa área tem sido assunto de investigação de pesquisadores envolvidos com computadores, Controle Automático, Teoria da Informação, Física Aplicada, Estatística, Psicologia, Fisiologia, Biologia, Medicina e Lingüística entre outras. A figura 1 extraída de Gonzaga, (1991) mostra a relação de diversas áreas do conhecimento com o Reconhecimento de Padrões. No sentido mais amplo padrões são os meios pelos quais pode-se interpretar o mundo. O termo padrão visual refere-se aos objetos do mundo real passíveis de serem representados no mundo artificial dos computadores, através de técnicas de Visão Artificial.

Em Verhagen (1975) e Uttal (1975) pode ser vista uma correlação de várias áreas associadas com a detecção de forma e Reconhecimento de Padrões.

Figura 1.1 - Áreas de atuação de Reconhecimento de Padrões e sua relação com as diversas subáreas do conhecimento (Gonzaga, 1991).

O problema básico no entendimento de como os seres humanos percebem a informação no constante bombardeamento do fluxo sensorial, consiste em um dos maiores problemas da Ciência Cognitiva, entretanto o entendimento da percepção é crítico para o desenvolvimento de uma teoria completa sobre o processamento humano da informação porque a percepção serve de interface entre os processos do mundo físico com o mundo mental. Os padrões de energia que atingem os receptores sensoriais devem ser transformados em representações que são canônicas com as propriedades dinâmicas e estruturais do mundo físico. Ao mesmo tempo essas representações devem ser compatíveis com os processos cognitivos envolvidos na percepção, compreensão e memória. As teorias da percepção entretanto devem considerar a física da luz e do som, a estrutura de objetos e linguagem, a Neurofisiologia do sistema sensorial, as limitações e habilidades da atenção, aprendizagem e memória, e as restrições computacionais do processamento paralelo e serial.

A Percepção não pode ser apenas entendida através de uma descrição completa dos estímulos físicos do código neuronal que é usado para representar o produto de uma transdução sensorial. O reconhecimento não ocorre nos órgão sensoriais. Similarmente, novos algoritmos para comparação de padrões, ou teorias de memória ou atenção, ou até mesmo novas teorias básicas de computação não poderão sozinhas resolver o problema. A solução para o entendimento da informação perceptiva humana dependerá de uma

abordagem inter-disciplinar que integre conhecimento científico sobre a psicologia cognitiva, computação, física, matemática e lingüística.

2 - MOTIVAÇÕES TEÓRICAS

Um problema fundamental em Visão Computacional é representado pela complexidade computacional de algumas tarefas como detectar um elemento alvo em uma imagem (Busca Visual) e o problema de achar uma correspondência entre a imagem e um conjunto de modelos (*matching*). Uma solução aparente para este problema seria considerar uma certa perda da informação e estruturar o processo de *matching* em um número maior de estágios. Por exemplo, construindo-se um número de representações de propósito geral da imagem tais como orientação e mapas de características de cor, e usar estes mapas associados à informações *top-down* para acessar os modelos em uma base de dados.

Um aspecto relacionado a busca visual e ao uso de uma representação distribuída para múltiplos objetos é a dificuldade de distinguir o objeto alvo entre outros elementos distrativos e a dificuldade de integrar (*binding problem*) corretamente as diferentes características extraídas nos mapas de características. Experimentos psicofísicos sugerem que pessoas utilizam a atenção visual encoberta para contornar esse problema (Ahmad, 1991). Na busca visual, dois aspectos do problema são importantes: integração de características e localização. A integração de características preocupa-se com a interferência entre as características de objetos diferentes quando é utilizada uma representação paralela. Similarmente, a interferência entre objetos dificulta a recuperação das localizações dos objetos individuais.

A partir de uma variedade de níveis hierárquicos (da retina até os estágios mais altos do córtex visual), os mecanismos de seleção descartam a maioria das informações a fim de concentrar as fontes limitadas de processamento para as regiões mais interessantes da imagem. Nos níveis mais altos além da análise de características, processos seletivos são necessários para garantir uma coerência comportamental (atenção para ação). Uma vez que a percepção é um meio que permite a interação (manipular, evitar, etc.) com objetos que compõem o ambiente, um grande número de ações são constantemente causadas e guiadas pela percepção do objeto. Cada uma dessas ações necessita da especificação de um número de parâmetros que podem ser vistos como controladores da identidade, posição e aparência

dos objetos. O processamento seletivo seria então necessário para isolar as informações que definem os parâmetros para as ações apropriadas. Se cada fluxo de processamento da informação gerado por algum estímulo visual potencialmente leva a uma ação, então existe a necessidade de mantê-los separados uns dos outros. Além do mais, essas informações devem ser coordenadas apropriadamente a de atender completamente a um estímulo, mas também deve permitir um desvio da atenção se novos eventos importantes acontecerem (Milanese, 1991).

A importância de selecionar a informação relevante da imagem é amplamente conhecida em Visão Computacional e é a base da recente “*Visão Ativa*”. Está claro agora que as suposições em que se baseiam os mecanismos atencionais, isto é, a possibilidade de definir corretamente o tamanho da região que contém um número limitado de objetos, não são irreais na maioria das aplicações práticas.

As vantagens computacionais de tais mecanismos vêm sendo explorada para solucionar os principais problemas fundamentais em Visão. Um exemplo onde o módulo de atenção visual é de grande importância é o controle de parâmetros de uma câmera na seleção de uma região de interesse na imagem. Esta região pode corresponder tanto a objetos estáticos que atraem a atenção ou ao contrário, aos objetos móveis que estão dentro do campo de visão.

Através da integração entre diferentes campos de pesquisa, muito progresso tem sido realizado. Pesquisadores em sistemas de Visão Artificial e em Psicologia, Biologia e Neurociência têm realizado um grande esforço para compreender a visão no homem e nos animais, no entanto ainda existe uma enorme distância entre nossa compreensão das funções visuais de alto nível e a organização neural de baixo nível. O modelamento das funções visuais, por exemplo em redes neurais pode contribuir para reduzir essa distância, providenciando um meio de implementá-las em um *hardware* neural de baixo nível (Arbib, 1989).

Na prática, somente é possível construir uma rede neural com um número fixo de elementos de entrada e de processadores internos. Os objetos de uma cena considerada podem estar afastados o suficiente uns dos outros, tal que seja impossível para o sistema “vê-los” simultaneamente. Dando-se como entrada a cena completa, um sistema neural não será capaz de processar toda a informação em paralelo a menos que exista um número

exponencial de unidades e conexões (Tsotos, 1988). A opção é processar uma pequena quantidade de informação em paralelo, focalizando sucessivamente em diferentes partes da cena (Desimone, 1989). Os seres humanos por exemplo, resolvem esse problema movendo seus olhos e focalizando a atenção em diferentes objetos em diferentes unidades de tempo (Hochberg, 1978).

Existem duas teorias complementares que foram propostas para explicar a atenção visual seletiva. A teoria baseada no objeto (Duncan, 1984) sugere que a atenção é sempre focalizada em um único objeto ou em um grupo coerente de características. A segunda teoria baseia-se no espaço e sustenta que a atenção é restrita somente a uma pequena região do espaço conhecida como campo atencional. Resultados experimentais mostram que objetos próximos do observador são percebidos mais rapidamente do que objetos distantes (Anderson, 1990). No plano frontal de visão, objetos mais próximos do centro do campo atencional são mais fáceis de reconhecer do que os que estão fora do seu centro (Eriksen & James, 1986). O trabalho de Treisman e co-autores (Treisman, 1986; Treisman & Gelade, 1980) demonstram a necessidade de localizar a atenção para a percepção visual.

Essas teorias sugerem o desenvolvimento de sistemas que focalizam sua atenção para reconhecimento em um componente da cena por vez, daí a proposta de um sistema de Atenção Visual Seletiva que decidirá sobre as prováveis regiões para a localização do objeto e deslocará para lá sua atenção, permitindo, através de um processo *top-down*, que o objeto seja reconhecido e sua posição localizada.

3 - OBJETIVOS

Existem diferentes aspectos que são identificados nas tarefas de reconhecimento. Este trabalho prevê o desenvolvimento de uma nova metodologia de atenção visual, baseada na Lógica *Fuzzy* e na Teoria dos Momentos, para operar em cenas simples, contendo um número limitado de objetos distribuídos aleatoriamente e abrange uma classe de tarefas envolvidas com objetos bidimensionais ou tridimensionais que possam ser projetados em duas dimensões mantendo entretanto suas características significativas. Como exemplos destes últimos, temos objetos cuja altura seja irrelevante quando comparada com a sua largura (por exemplo arruelas, porcas, parafusos, chaves, talheres, pratos, etc.). Os objetos também são considerados segmentados, não oclusos, e que consequentemente podem ser

analisadas utilizando técnicas baseadas em características globais. As mesmas características podem ser extraídas das cenas não-segmentadas, fornecendo uma informação de “textura”, ou seja da distribuição dos objetos dentro da cena.

Os objetivos fundamentais deste trabalho são:

- Contribuir com o desenvolvimento de visão que possam ser adaptados á diferentes aplicações, principalmente na área de inspeção automática. O método proposto objetiva encontrar na imagem de entrada os objetos que contenham informações mais “interessantes¹” e “importantes”.
- Contribuir para uma melhor compreensão do sistema visual humano através da modelagem de fenômenos de percepção de alto nível, como a Atenção Visual Seletiva.

Neste trabalho abordamos os seguintes aspectos:

- Uso de uma abordagem *fuzzy* para busca e reconhecimento de padrões visuais, baseada na integração de características *bottom-up* (provenientes dos objetos) e *top-down* (dependentes da tarefa), a fim de se detectar regiões de interesse onde se possa restringir processamentos mais complexos.
- Uso de uma abordagem de reconhecimento baseada em características independentes, extraídas dos objetos dentro da cena, diminuindo e controlando a dimensionalidade da informação, mantendo porém as informações essenciais que permitam sua discriminação.
- A adição de um sistema de reconhecimento de padrões invariante a transformações geométricas (através do uso de Invariantes de Momento IM) e capaz de trabalhar com múltiplos objetos (através de Mapas de Características) em uma imagem tem a finalidade de aumentar a flexibilidade dos sistemas invariantes de propósitos e aplicações mais gerais.
- Avaliação do sistema quanto ao uso em inspeção visual automática voltada para aplicações industriais.

4 – DESCRIÇÃO GERAL DO TRABALHO

Pode-se definir a Atenção Visual como a capacidade dos sistemas visuais biológicos em detectar rapidamente partes interessantes da imagem da retina, a fim de reduzir a

¹ O termo “interessante” utilizado aqui refere-se às características salientes de conspicuidade detectadas no subsistema *bottom-up*. Tais características podem revelar alguma medida de irregularidade (diferenças) entre os padrões presentes na cena. Características interessantes podem também referir-se a qualquer atributo do alvo, necessário para realizar uma tarefa específica, no subsistema *top-down*.

quantidade de dados para tarefas mais complexas. Jonhston & Dark (1986) definem atenção visual como o processamento diferencial de várias fontes simultâneas de informação. Busca Visual, em nosso trabalho refere-se à tarefa de se detectar dentro do campo visual objetos alvos específicos (*targets*) utilizando-se as informações de formato disponíveis (*bottom-up*) e o conhecimento prévio sobre a tarefa e sobre as características de formato que definem o objeto (*top-down*) a fim de se facilitar suas detecções. A Atenção Visual pode ser vista como uma “cola” que une fontes de informações internas (memória e conhecimento) e fontes de informação externas (objetos e eventos). Sem a atenção, características não podem ser relacionadas umas com as outras.

A análise visual da forma e de suas relações espaciais exerce um papel importante no planejamento de ações e na manipulação de objetos. Esta análise não ocorre no processamento visual primário (*early processing*) mas possui grande importância nos estágios posteriores (nível intermediário – *intermediate level*). É impossível extrair todas as propriedades relevantes de formato e todas as relações espaciais na imagem, de uma maneira *bottom-up* uniforme (*low-level*). Diferentes tarefas necessitam da extração de diferentes propriedades de formato e suas relações (Hildreth & Ullman, 1990). Segundo Ullman, (1984) o sistema visual possui um conjunto de operações básicas construídas que podem ser usadas para se construir diferentes rotinas visuais e para extrair uma grande variedade de propriedades de formato e suas relações. Isto implica em uma estrutura espacial e temporal do processo. Uma certa operação é aplicada em um certo ponto do campo visual e dependendo dos resultados, uma outra operação é aplicada em outro ponto.

O reconhecimento do objeto necessita algum tipo de comparação (*matching*) entre a representação interna do objeto armazenado na memória e uma representação similar gerada pelo objeto na imagem. Este processo está essencialmente relacionado a organização, recuperação da memória, expectativa, raciocínio, portanto é diferente dos níveis baixo e intermediário. Por definição processo de alto nível é todo aquele que produz ou utiliza imagens mentais enquanto que processo de baixo nível é aquele que utiliza exclusivamente as propriedades (características) extraídas dos estímulos presentes na imagem da retina (por ex. segmentação).

Hildreth & Ullman, (1990) definiram como deslocamento (*shifting*) a operação de mover o foco atencional de processamento para um ponto requisitado e indexação (*indexing*)

um tipo particular de *shifting* onde o processamento é deslocado para o ponto mais *saliente* (*relevante*) do campo visual.

Na maioria dos trabalhos, a localização do próximo foco é computada utilizando-se um mapa retinotópico que codifica a importância de cada posição em cada estado de interpretação. Dois tipos principais de processos *bottom-up* e *top-down* contribuem para quantificar a importância em cada localização. Objetos que diferem muito de seu ambiente, por exemplo em cor, orientação, ou tamanho, atraem a atenção de uma maneira *bottom-up*. Na atenção *top-down*, os altos níveis cognitivos do cérebro influenciam o sistema atencional a selecionar em favor de uma característica em particular ou em favor de uma combinação de características. A primeira classe é puramente dirigida pelos dados (*data-driven*) enquanto que a segunda inclui restrições sobre a tarefa e depende do conhecimento já obtido da imagem.

Uma abordagem plausível para o reconhecimento humano assume que não armazenamos e comparamos todos os diferentes exemplos de padrões, e sim produzimos algum tipo de representação interna armazenada na memória e que permanece *invariante* em relação a um grande número de vistas.

A abordagem empregada neste trabalho integra características de formato baseadas na extração de Momentos Centrais (e seus atributos) e os Invariantes de Momento (IMs) em um sistema que faz reconhecimento de objetos e análise de cenas invariantes às transformações geométricas e tolerante a um certo grau de ruído. Os Momentos Centrais (MC) e seus diversos atributos de formato, bem como os Invariantes de Momento (IM) transformam a imagem de entrada em diversos mapas de características de orientação, simetria, tamanho, alongação, invariância à rotação, “*textura do formato*” referente ao objeto (contorno, simetria), além de fornecer informações sobre a cena quanto a distribuição dos objetos. Momentos foram escolhidos devido a simplicidade de implementação e por fornecerem uma interpretação “física” de formato baseado na distribuição da massa dos objetos nos eixos x e y. Os Invariantes de Momento perdem esta informação de distribuição dos objetos mas apresentam invariância à rotação e também podem ser muito úteis no reconhecimento de objetos pois fornecem uma característica de formato. Os Invariantes Momentos têm sido amplamente utilizados em imagens segmentadas para Reconhecimento de Padrões aliados à Redes Neurais Artificiais (RNA) como classificadores, mas pouca

ênfase tem sido dada aos atributos de formato fornecidos através da aplicação de Momentos Centrais tanto em objetos, quanto em cenas.

A percepção de objetos que podem ser reconhecidos em diferentes pontos de vistas ou em diferentes posições no espaço deve-se também a divisão do caminho retinal em dois caminhos visuais. Mishikin *et al.*, (1983) mostrou que o caminho que atinge a área parietal é crucial para localizar objetos (“*Where*”), enquanto que outros que alcançam a área temporal é crucial para o reconhecimento de forma (“*What*”). Neurônios do caminho temporal possuem campos receptivos muito grandes, podendo atingir até mesmo todo o campo visual e isto faz com que o reconhecimento no caminho “*What*” seja independente da posição do objeto na retina.

Devido a esta divisão da informação, também é necessário combiná-la a fim de que seja possível a localização e identificação do objeto ao mesmo tempo. Esta integração entre a informação disponível (*bottom-up*) e o conhecimento (*top-down*) é utilizada para otimizar a busca de objetos dentro do campo visual. As informações *bottom-up* são extraídas em paralelo e a *importância (relevância)* da informação disponível é dada por exemplo através de uma medida de *conspicuidade (diferença)* ou através do *conhecimento prévio* sobre o objeto ou sobre quais características melhor o representa (controle *top-down*).

De acordo com Levine (1985), E. Mach foi quem iniciou o estudo moderno da forma. Ele investigou quais características afetam a nossa percepção da forma e sugeriu que a forma é uma entidade processada independentemente dos outros atributos de baixo nível. Max Wertheimer publicou em 1912 as bases da teoria Gestalt, contribuindo com um progresso real para o entendimento humano da organização da percepção.

Segundo a psicologia Gestalt, objetos parecidos tendem a formar um agrupamento perceptivo enquanto que objetos muito diferentes entre si ou com atributos diferentes tais como cor, brilho, contraste, textura, formato, simetria, orientação, alongação, distorção, distribuição e posicionamento dos objetos na cena (agrupamento), movimento, etc., despertariam a atenção. Esta *diferença (“distância”)* entre os atributos dos objetos é chamada de *conspicuidade (saliência)* e significa “*aquilo que se ressalta por ser diferente*”. A conspicuidade mede quão diferente um objeto é do outro e em alguns casos pode ser útil utilizar esta abordagem para detectar rapidamente objetos alvos na imagem.

A teoria Gestalt estabelece também que devemos prestar atenção ao estímulo global do padrão, mais do que em suas partes constituintes e que os mecanismos da percepção da forma devem ser holísticos. A Gestalt apresenta também um conjunto de leis que governam a maneira com que agrupamos os estímulos juntos e a maneira com que separamos a figura do fundo². De acordo com a teoria Gestalt, inicialmente registramos objetos e relacionamentos unitários e apenas mais tarde, se necessário, analisamos esses objetos em suas partes constituintes. A psicologia cognitiva considera a percepção como um processo mental e estabelece que a percepção global é tomada em termos de seus atributos. Apesar dessas idéias parecerem contraditórias, pesquisadores perceberam que o contexto global (*Whole*) depende de operações complexas as quais uma pessoa não possui acesso consciente. Como dito por Haber & Wilkinson (1982), todas as partes de um objeto são percebidas juntas e todos os objetos estão associados uns aos outros com relacionamentos do tipo: perto, longe, atrás, junto.

Recentes progressos em Neurociência vêm clareando mais sobre a estrutura e organização dos centros cerebrais que cooperam com as funções visuais. Especialização, modularidade e organização hierárquica parecem ser fundamentos importantes para entender a operação desse sistema paralelo extremamente complexo. Existem muitas razões para acreditar que a visão aplica analisadores especializados para decompor estímulos em partes e propriedades. Operações extras são necessárias para especificamente recombina-las dentro de um “todo” correto. Essa evidência é em parte anatômica e fisiológica. Uma área preocupa-se com a orientação de linhas e bordas, outra área com percepção de cor e assim por diante. Apenas após o processamento nessas áreas, os dados alcançam áreas que discriminam objetos complexos.

O modelo proposto está estruturado em três estágios principais, conforme ilustra a figura 1.2.

² O agrupamento perceptivo pode ser definido como a capacidade de agrupar ou organizar elementos menores em grandes configurações. A característica de forma baseada em momentos podem produzir comportamento análogo.

Figura 1.2 – Diagrama em blocos do Modelo Proposto

O primeiro estágio consiste em segmentar os objetos e extrair características globais dos mesmos baseadas principalmente na teoria dos momentos, tais como média de nível de cinza, tamanho, orientação, formato e distância. Os atributos utilizados para guiar a atenção foram escolhidos baseando-se nas evidências obtidas no estudo da atenção humana. A abordagem apresentada baseia-se na decomposição da imagem de entrada em um conjunto de mapas de características independentes. Cada mapa representa o valor de um certo atributo computado para um conjunto de objetos. Objetos relevantes são detectados se as primitivas correspondentes apresentarem valores de características que diferem fortemente dos demais objetos presentes na cena. Através da comparação de um objeto com os outros presentes na cena, características *bottom-up* de conspicuidade são usadas para guiar a atenção ao objeto mais diferente. Os conjuntos *fuzzy* (*fuzzy sets*) do sistema proposto baseiam-se nestas características e foram escolhidos devido a sua influência em caracterizar os processos *bottom-up* e *top-down* da atenção. O segundo estágio consiste de um subsistema *fuzzy top-down* que combina diferentes características de acordo com a relevância das mesmas em diferentes tarefas. Finalmente o terceiro estágio consiste de um subsistema *fuzzy* que integra as informações obtidas dos subsistemas anteriores e fornece um índice geral de saliência indicando a provável localização do objeto a ser reconhecido. A nova abordagem foi testada com objetos geométricos levando-se em consideração as características que atraem a atenção dos seres humanos. Através do uso da lógica fuzzy foi possível extrair o conhecimento humano para estabelecer as funções de pertinência (*membership functions*) que determinam o grau de pertinência de uma determinada variável

INSTITUTO DE PESQUISA EM CIÊNCIAS

dentro do correspondente *fuzzy set*. O subsistema *fuzzy* foi elaborado com 243 regras que combinam os cinco conjuntos *fuzzy* de entrada (conspicuidade de *gray*, conspicuidade de tamanho, conspicuidade de orientação, conspicuidade de forma e conspicuidade de distância) com suas respectivas funções de pertinência (*small*, *medium* e *large*). O conjunto *fuzzy* de saída fornece um índice de saliência *bottom-up*. O subsistema *fuzzy top-down* é orientado à tarefa e foi construído com 252 regras e tem o objetivo de fornecer um índice de saliência *top-down*. No trabalho criou-se uma base de regras para armazenar diferentes tarefas tais como “pegar o menor alicate”, “pegar a chave de cano mais escura e mais próxima” e assim por diante

5 - ORGANIZAÇÃO DA DISSERTAÇÃO

A importância de se construir uma metodologia que baseia-se nos mecanismos da Atenção Visual em aplicações de Visão de Máquinas foi apresentada nas seções anteriores e serviu também como fonte de inspiração para o desenvolvimento deste trabalho.

Este trabalho, além desta introdução, está estruturada em duas partes. A primeira parte envolve três capítulos e tem o objetivo de levantar e relacionar os vários fundamentos relativos aos aspectos psicológicos e computacionais envolvidos na Atenção Visual e da Percepção³ de Forma.

O segundo capítulo relaciona os principais aspectos psicológicos⁴ e psicofísicos⁵ relativos à percepção de objetos e também apresenta uma seleção de algumas teorias de

³ Percepção: consciência, discernimento (Michaelis).

Perceber: Adquirir conhecimento de (aprender), por meios dos sentidos (Aurélio).

⁴ Psicologia – Ciência dos fenômenos psíquicos e do comportamento.

Fenômenos psíquicos – conjunto de processos conscientes ou inconscientes de um indivíduo.

A Atenção (ao nível da percepção) é um fenômeno psíquico relacionada com percepção consciente.

⁵ Psicofísica – Estudo do relacionamento entre estímulos físicos e a resposta perceptiva do observador ao estímulo. Estudo das relações funcionais entre a mente e os fenômenos físicos. Assim a Psicofísica relaciona os primeiros estágios do processo perceptivo (estímulo) e o estágio final (reconhecimento), excluindo os passos fisiológicos⁷ entre eles. Apesar disso, resultados psicofísicos muitas vezes tornam-se pontos de partida para estudos fisiológicos.

Psicopictórica – subcampo da Psicofísica que lida com imagens naturais como estímulos pictóricos. Mais especificamente baseia-se na proposição de que em uma imagem, a informação referente ao observador humano pode ser caracterizada e analisada em termos das propriedades percebidas do objeto.

Percepção Humana da Forma. Procurou-se relacionar também a Psicologia Gestalt com a Atenção Visual.

O terceiro capítulo apresenta os principais conceitos relacionados a Atenção Visual humana, como definições, motivos para sua existência, tipos, busca visual, influência da memória e expectativa, desenvolvimento na infância, entre outros. Neste capítulo são apresentados também vários modelos para a Atenção Visual Humana e no quarto capítulo são revistos também diversos modelos para Atenção Visual em Máquinas realizados pelos cientistas nos últimos anos.

A segunda parte deste trabalho descreve os detalhes do modelo desenvolvido em relação a diversos aspectos de projeto tais com aquisição da imagem, escolha das características, estratégia de busca visual de objetos, construção de mapas de características e de mapas de conspicuidade que representam diferentes graus de interesse em cada localização na imagem, construção da base de conhecimento baseado na lógica *fuzzy* (base de regras e de funções de pertinência) e a definição do processo de integração que agrega o índice de atenção *bottom-up* e o índice de atenção *top-down* e determina um índice de atenção global para os objetos presentes na cena, de acordo com a disposição e características dos objetos e a relevância da tarefa. Por último são apresentadas algumas conclusões e comparações em relação a algumas abordagens existentes, ressaltando as vantagens e limitações da metodologia proposta.

Resultados provenientes da Neurobiologia, Neurofisiologia⁶ e da Psicologia Cognitiva são de particular interesse, uma vez que eles estudam diferentes aspectos funcionais da Atenção Visual, bem como suas relações com outras funções motoras e visuais e podem contribuir para o desenvolvimento de novas técnicas de análise de forma. Por isso, a fim de complementar alguns conceitos apresentados no capítulo 3, o apêndice 1 enfatiza os aspectos fisiológicos⁷ relativos ao reconhecimento de objetos pelo Sistema Visual Humano. O apêndice 2 é dedicado a percepção computacional de formas através da

⁶ Distinção feita entre Psicologia e Neurofisiologia: "...os objetos de atenção da Psicologia⁴ são pensamentos, percepções, sensações e outros aspectos da experiência consciente. "os objetos de atenção da Neurofisiologia são os processos que são realizados por tecidos especializados e células do sistema nervoso".

⁷ Fisiologia – Parte da Biologia que investiga as funções orgânicas, processos ou atividades vitais como visão, respiração, etc. Estuda a relação do estímulo visual com a resposta neural
Psicofisiologia: Estudo científico das relações entre a atividade fisiológica e o psiquismo (mente) .

teoria dos momentos, que constitui a base para as características utilizadas neste trabalho. Várias características baseadas na interpretação geométrica dos Momentos são mostradas e revistas. No apêndice 3 são apresentadas as noções básicas da teoria dos conjuntos *fuzzy* e seus principais modelos de inferência. O apêndice 4 mostra como foram criadas as tabelas de regras utilizadas para a realização de tarefas específicas pelo sistema *top-down*. Todos os arquivos pertencentes a este trabalho estão disponíveis em um CD-ROM que acompanha esta tese. No apêndice 5 (disponível apenas no CD-ROM) podem ser vistas as imagens utilizadas nos testes psicofísicos para a determinação das funções de pertinência dos conjuntos *fuzzy* de entrada dos subsistemas *bottom-up* e *top-down*. Em toda a dissertação, adotou-se a fonte itálico para os termos estrangeiros.

CAPÍTULO 2

PSICOLOGIA DA FORMA

O todo é igual a soma de suas partes (Euclides)

O todo é maior que a soma de suas parte (Max Wertheimer)

1 - INTRODUÇÃO

A inteligência¹ é um fenômeno natural que foi desenvolvida pelos seres vivos a fim de que os mesmos pudessem se adaptar às mudanças ocorridas no meio ambiente resultantes de suas próprias ações e também devido a agentes externos e processos naturais, e pode ser abordada sobre duas perspectivas: cognição e percepção. A cognição diz respeito à natureza do raciocínio humano enquanto que a percepção diz respeito a aplicação desses mecanismos para modelar e entender o mundo externo, baseando-se na informação sensoriada.

O Sistema Visual Humano (SVH) é um dos mais sofisticados e versáteis da natureza e muita pesquisa tem sido dedicada com a intenção de se construir um modelo para sistemas artificiais de percepção visual e cognição. As primeiras pesquisas na área de sistemas artificiais para análise de forma estavam isoladas da área de percepção visual humana, em oposição às pesquisas recentes que baseiam-se mais nos resultados provenientes do estudo do SVH e da percepção.

Tudo o que vemos, vemos pela primeira vez. Mesmo que algumas partes da cena corresponda á objetos conhecidos previamente, quase sempre nunca vemos os mesmos objetos, nas mesmas configurações, sob as mesmas condições de luz e na mesma perspectiva no espaço. A menos que possamos decompor a cena em entidades reconhecíveis, coerentes e independentes, a complexidade das cenas naturais tornaria a visão humana impossível. Como podemos analisar (dividir) a cena em componentes independentes sem antes conhecer o que pode estar presente? Se procurássemos apenas por objetos bem conhecidos, teríamos

¹ A inteligência pode ser definida como a faculdade de aprender (cognição), compreender (percepção), interpretar. Capacidade de adaptar-se facilmente, agudeza (acuidade) ou seja, capacidade de discriminar estímulos sensoriais.

Cognição: atividade mental (Goldstein, 1989).

que determinar exaustivamente se cada objeto estava presente na cena em cada localização possível do campo visual. Entretanto existem provavelmente milhares de objetos que podem aparecer em uma variedade infinita de configurações e orientações, dessa maneira uma comparação exaustiva com todos os modelos armazenados não constitui uma explicação razoável da percepção humana. A figura 2.1 extraída de Levine, (1985) ilustra o problema.

Figura 2.1 – Diversas representações para o objeto “cadeira” (Levine, 1985).

Percepção é uma experiência constante e individual, *modelada* pelo processamento neuronal e que pode permanecer constante na mudança de estimulação, conforme será visto nos próximos capítulos. *“Percepção não é um armazenamento passivo do ambiente pelos órgãos sensoriais. A Percepção consiste de uma construção ativa por meios de quais dados estão sendo selecionados, analisados e integrados com propriedades não percebidas diretamente e apenas gerando hipóteses, deduzindo ou antecipando, de acordo com a informação disponível e capacidade intelectual.”* Gaetano Kanizsa apud Sarkar & Boyer (1993).

A maioria das teorias da percepção focaliza uma pequena área dentro do grande campo da percepção. Assim existem teorias para a percepção da cor, percepção de intensidades, teorias de ilusões visuais e teorias de percepção de formas². Neste trabalho

² Segundo Pomerantz (1986) forma é uma entidade psicológica. Um objeto físico possui as mesmas características e estrutura, independente de como (ou se) é percebido, entretanto um objeto camuflado em seu fundo não constitui uma forma a menos que seja percebido como tal.

será dada uma ênfase maior para as teorias de percepção de forma principalmente para a teoria Gestalt que aborda a percepção global e a organização perceptiva e também para as teorias de Treisman relacionadas à atenção e busca visual e a de Ullman relacionada à construção de rotinas visuais.

Psicólogos interessados na organização da percepção questionam como estímulos discretos podem ser vistos como um único estímulo. A determinação das unidades básicas de análise para estímulos multidimensionais ou estruturados hierarquicamente é basicamente o problema da percepção da forma, porque formas são estruturas multidimensionais e hierarquizadas. Como os psicólogos Gestalt demonstraram, existem algumas características elusivas emergentes que não são medidas facilmente e que desaparecem quando analisadas suas partes constituintes. Tem havido um ressurgimento do interesse nos fenômenos básicos da psicologia Gestalt, tais como agrupamento e segregação figura-fundo e o efeito do contexto na percepção.

2 – TEORIAS DE PERCEPÇÃO HUMANA DE FORMA

No que diz respeito ao processo de reconhecimento de objetos, Ernest Mach (1914) foi quem providenciou o ímpeto para o moderno estudo das formas. Ele investigou quais características caracterizavam nossa percepção da forma e sugeriu que a forma era uma entidade que de algum modo era processada de maneira independente dos outros atributos de baixo nível.

Seres humanos e poucos outros organismos vivem em um mundo de cores e formas. Percebemos a nós mesmos como nos movimentando em um ambiente estacionário. Podemos reconhecer e perceber objetos comuns como tendo formas esperadas, até mesmo quando vistos em diferentes distâncias, orientações e sobre diferentes condições de iluminação. Assim, se olharmos obliquamente um círculo desenhado em uma superfície plana, podemos ver a figura esperada do círculo, mesmo quando a imagem projetada em nossa retina for uma elipse. Podemos ajustar mais de 10 vezes a magnitude da intensidade de luz se não nos preocupamos com sombras e oclusões parciais.

É óbvio que a percepção não é um resultado inevitável de um conjunto de padrões de estímulo, mas uma interpretação de dados sensoriais baseados na experiência passada

dos organismos e de seus ancestrais. Enquanto os sentidos não nos fornecem um modelo fiel do mundo, eles podem entretanto fornecer evidências para hipóteses sobre o ambiente que nos rodeia. A maneira mais concisa de resumir nossa capacidade visual é que ela parece operar de uma maneira impecável, espontânea e sem surpresas.

Nos níveis mais altos de sua realização, o reconhecimento de formas envolve a habilidade de ignorar variações em tamanho, brilho, posição e orientação. Alguns animais superiores tais como ratos e polvos, não possuem habilidade para lidar com variações de 90 ou 180°. Uma possibilidade para isso é que enquanto o sistema visual humano pode geralmente decompor o campo visual em sub-unidades, a maioria dos outros organismos não possuem essa habilidade. Se um sistema visual não pode extrair e manipular partes de uma imagem, as operações de normalização necessárias para uma visão de padrões robusta torna-se complexa e impossível de ser implementada. Mesmo organismos mais simples possuem a habilidade para reconhecer padrões. Abelhas, formigas e vespas, por exemplo, utilizam marcas visuais para voltarem para seus ninhos. Abelhas podem reconhecer sua colmeia através de marcas coloridas em sua entrada e pode determinar se as cores possuem uma disposição espacial apropriada quando se aproximam da colmeia em uma direção arbitrária. Entretanto abelhas não podem distinguir entre formas geométricas mas respondem em tais casos se as figuras forem ramificadas ou divididas em partes (Levine, 1985).

Várias escolas de psicologia se empenharam em entender e descrever os mecanismos do comportamento em geral e nos aspectos específicos da percepção visual. Algumas teorias de percepção visual de forma podem ser vistos em Hake, (1966), Zusne (1970 a,b), Goldstein, (1989). Algumas teorias mais recentes podem ser vistas em Loncaric, (1998). Um tutorial em Cognição Visual com ênfase no reconhecimento de forma foi escrito por Pinker (1984).

2.1 – Abordagem Estruturalista

Sensação e Percepção são coisas diferentes. Em 1785 o filósofo Thomas Reid apud Goldstein, (1989) propôs que a diferença crucial entre elas é que percepção sempre se refere á objetos externos, enquanto que sensações referem-se à experiências dentro de uma pessoas, que não estão ligadas á objetos externos. Quando olhamos para o céu, por exemplo, e vemos a luz de uma estrela, estamos *sentindo* um pequeno ponto de brilho mas logo que vemos esse ponto relacionado a um objeto, tal como uma estrela, estamos *percebendo* uma

estrela. A idéia de que percepção refere-se à objetos externos levou alguns pesquisadores a proporem que crianças no início da vida experimentam apenas sensações e encontram percepção após terem adquirido alguma experiência.

Uma abordagem diferente para distinguir entre sensação e percepção foi feita pelos Estruturalistas no início do século que propuseram que sensações são blocos elementares de percepções. De acordo com essa teoria, percepções são criadas pela adição de sensações. A importância dessa abordagem é principalmente histórica pois atualmente já não é aceita. Suas idéias estimularam o nascimento da psicologia Gestalt, que ainda prospera até hoje.

2.2 – Abordagem Gestalt - 1910

No fim do século XIX surgiu a escola da “Psicologia da Forma” (Gestalt) que combateu a psicologia associacionista que sustentava que a percepção global se reduz a uma soma. Gestalt é uma palavra alemã que significa a formação de um *todo* (estrutura global) organizado e significativo. A palavra Gestalt “tem o significado de uma entidade concreta, individual e característica, que existe como algo *destacado*³ e que tem uma forma ou configuração como um de seus atributos” (Köhler, 1929). Portanto uma Gestalt é uma organização; a organização é o processo que leva a uma Gestalt. No processo de organização “o que acontece a uma parte do todo é determinado por leis intrínsecas inerentes á esse *todo* (*whole*) ” (Wertheimer, 1925). Dizer que um processo ou o produto de um processo é uma Gestalt significa que ele não pode ser explicado pelo mero caos, a mera combinação de causas essencialmente desconexas mas que sua essência é a razão de sua existência (Koffka, 1935).

Se vemos um disco sobre uma mesa, dizemos que é um disco mesmo que o vejamos obliquamente, no entanto deveríamos dizer que se trata de um “oval”. A psicologia associacionista afirma que essa impressão está sujeita a uma transformação devido à nossa experiência em manusear discos, portanto não se trata de uma experiência verdadeira, pura e simples, e sim de uma série de sensações genuínas e de impressões vividas anteriormente. A figura 2 abaixo, extraída de Koffka, (1937) ilustra um caso semelhante ao exemplo acima. Segundo os assosacionistas vemos a figura 2.2 (a) como um decágono com duas linhas em seu interior porque temos visto essa figura mais freqüentemente do que as quatro outras

³ - Deseja-se mostrar que, no sentido da percepção (consciência) a Gestalt, que tem uma relação direta com a atenção visual (consciente), como algo que se destaca (saliente) e que tem uma estrutura.

figuras pequenas e irregulares. Wertheimer, (1923) apud Koffka, (1935) replicou à essa objeção construindo figuras como a 2.2 (b), na qual estão escondidas as iniciais de seu nome M e W.

Figura 2.2 - Exemplos de figuras vistas preferencialmente como globais (Koffka, 1935).

A psicologia Gestalt afirma exatamente o contrário. Segundo ela o estímulo local é praticamente irrelevante e o que se percebe depende unicamente da condição global do organismo, isto é, da estimulação global do órgão sensitivo interessado, ou melhor, da estimulação de muitos ou de todos os órgãos. A teoria Gestalt estabelece que devemos prestar atenção ao estímulo global do padrão, mais do que em suas partes constituintes e que os mecanismos da percepção da forma devem ser holísticos. Segundo a teoria Gestalt, inicialmente registramos objetos e relacionamentos unitários e apenas mais tarde, se necessário, analisaremos estes objetos em suas partes ou propriedades constituintes (Treisman & Gelade, 1980).

Para nós aquele disco é um disco independente do ângulo do qual o observamos. Existe portanto um fenômeno de constância da forma, das dimensões e da cor. Por exemplo, quando deslocamos um disco diante dos nossos olhos à uma distância de 1 para 10, as dimensões na retina variam na mesma proporção e no entanto não temos absolutamente a sensação de que o disco tenha diminuído dez vezes. Da mesma maneira, se o disco é branco, continuamos a julgá-lo branco mesmo que o ambiente esteja mais escuro. Se observamos um objeto que se desloca com velocidade uniforme à distâncias diferentes, a velocidade parece-nos mais ou menos a mesma, ao passo que na imagem da retina a variação com a distância é bem grande. O martelo não nos parece mais pesado se o levantamos com um único dedo em vez de o fazermos com a mão, ao passo que de acordo com os princípios associacionistas deveria parecer mais pesado, visto que a superfície de estímulo é menor. Por sua vez a redundância de conexões e mecanismos no sistema nervoso dos animais é tão grande que assegura uma margem de segurança e que certamente não favorece as teorias atomísticas associativas e a presumida correspondência de 1:1 entre estímulos e zonas reflexas (Zingales, 1978).

Portanto, segundo a Gestalt percebemos impulsivamente os objetos, conforme uma totalidade inerentes aos mesmos, ou melhor as constituímos como entidades fechadas, unicamente em decorrência de certos fatores e em pequena parte devido á uma experiência empírica anterior. Sem dúvida a aprendizagem não deve ser excluída mas considerada como uma extensão do nosso sistema primitivo de percepção.

A percepção visual pode falhar pela falta de estímulo mas também pela falta da informação de estímulo. Em ambientes homogeneamente escuros, a visão falha pela falta de estímulo e em um ambiente homogeneamente branco, por exemplo, a visão falha por falta de informação de estímulo. Neste caso, há estimulação da retina pela luz mas sem nenhuma ativação dos olhos pela informação do estímulo. A distinção entre estimulação para os receptores e informação de estímulo para o sistema visual é crucial (Gibson, 1986).

Segundo os psicólogos da escola Gestalt, Wertheimer, Köehler e Koffka, a percepção do ambiente baseia-se fundamentalmente na percepção de forma. O ambiente completamente homogêneo não pode ser percebido; parece vago, indefinido, sem localização no espaço. A percepção só ocorre na medida que surge alguma forma ou padrão e na medida em que são discriminados em virtude do contorno que os separa de seu fundo (segmentação figura-fundo), antes que isso ocorra não há percepção de forma. Entretanto, alguns estudos mostraram que as crianças podem ser menos capazes do que os adultos para diferenciar a figura do fundo (Gollin, 1956).

2.2.1 - Leis de organização

A Gestalt apresenta também um conjunto de leis que governam a maneira com que agrupamos (integramos) os estímulos juntos e a maneira com que separamos a figura do fundo. Apesar das leis descreverem bem alguns exemplos em que são aplicadas, sua operação não é sempre tão direta. Um problema consiste na aplicação da lei da simplicidade. Não é tão simples responder se uma figura é mais simples que a outra⁴. Algumas medidas são facilmente determinadas como luminosidade e proximidade. A figura 2.3 abaixo extraída de Goldstein (1989) ilustra o problema de se definir simplicidade. Muitas figuras podem ser vistas de mais de uma maneira. A forma ilustrada abaixo é percebida inicialmente como uma estrela de seis pontas bidimensional, entretanto várias alternativas perceptivas

⁴ Um interessante trabalho no sentido de definir simplicidade da forma e também na maneira em que extraímos informações significativas na imagem pode ser visto em Attneave (1957).

são possíveis, como mostrado na figura 2.3 (b) e 2.3 (c). Apesar da lei da simplicidade prever que a forma mais simples da percepção é ver a forma global bidimensional e simétrica da estrela, pessoas frequentemente vêem percepções mais complexas. A lei da simplicidade não fornece informações suficientes sobre o que é simplicidade para nós.

Figura 2.3 - Formas que podem ser vistas de diversas maneiras (Goldstein, 1989).

A organização da percepção, ou seja, o agrupamento de partes em unidades maiores é essencial para o mundo fazer sentido. Quando vemos uma paisagem campestre o processo de organização separa as árvores dos carneiros e vemos cada árvore como uma entidade diferente. Similarmente quando estamos conversando com alguém em um ambiente ruidoso, o processo de organização faz com que consigamos diferenciar a voz da pessoa dos outros sons que simultaneamente atingem nossos ouvidos.

Quando um borrão de tinta cai no papel ele é visto em primeiro lugar como uma *unidade segregada* do resto do papel e em segundo lugar tem uma configuração. Ele se torna segregado porque tem uma coloração diferente da circunvizinhança. As áreas que recebem a mesma estimulação homogênea organizarão partes unitárias do campo, segregadas das outras pela diferença entre estimulações. Segundo Koffka, (1937) a igualdade de estimulação produz forças de coesão, a desigualdade de estimulação produz forças de segregação, desde que a desigualdade acarrete uma mudança abrupta (*primeira lei de formação unitária e segregação*).

O sistema visual humano sugere que um conjunto de critérios genéricos, aplicado independentemente do conteúdo da cena, faz parte do procedimento descoberto pela natureza para dividir o campo visual. Descontinuidades nas propriedades das cenas, como por exemplo, distância, texturas e movimento, são as pistas mais utilizadas de onde devem ser inseridas as divisões. Psicólogos têm pesquisado as leis que regem a organização da percepção humana. Um dos primeiros trabalhos e mais aceito intuitivamente foi o conjunto de leis propostas por Max Wertheimer em 1912. Wertheimer, (1912) publicou as bases da teoria Gestalt, contribuindo com um progresso real no conhecimento humano sobre a

organização da percepção. Ele propôs um conjunto de leis de organização, as quais eram supostas estarem envolvidas na percepção de formas, bem como na segmentação das mesmas em relação ao fundo. A figura 2.4 abaixo ilustra o problema de como algumas formas brancas e pretas tornam-se perceptivamente organizadas na figura de um dalmata.

Figura 2.4 - Formas pretas e brancas que se tornam perceptivamente organizadas em um dalmata (Marr, 1982).

As leis da organização são um conjunto de leis que descrevem como nossa percepção é afetada dada certas condições de estímulos. Algumas das leis⁵ mais importantes que determinam em nós a organização do campo visual são mostradas a seguir (Levine, 1985; Goldstein, 1989, Fischler & Firschein, 1987):

- **Lei da vizinhança ou da proximidade** – Elementos de estímulos que estão próximos tendem a ser percebidos como um grupo. As partes de um todo são vistas segundo unidades reguladas pela distância mínima (ver Figura 2.5). Percebemos os pontos próximos e os constituímos como unidades.

Figura 2.5 - Ilustração da Lei da vizinhança.

- **Lei da similaridade, familiaridade ou da igualdade** – Estímulos similares tais como contraste, cor e formas, tendem a ser agrupados juntos; essa tendência pode dominar o

⁵ Helson, (1933) contou 114 leis de organização Gestalt.

agrupamento devido a proximidade. A percepção desses grupos não é fundamental para o reconhecimento de formas, mas é fundamental para discriminação de texturas. Na figura 2.6 verifica-se a tendência de agrupar os elementos semelhantes entre si. Percebemos como unidades os círculos cheios de um lado e os vazios do outro, ou as faixas de igual espessura.

Figura 2.6 - Ilustração da Lei da igualdade

- **Lei da compactidade (fechamento)** – Áreas com contornos fechados tendem a serem vistos mais como figuras do que áreas com contornos abertos. Estímulos tendem a ser agrupados em figuras completas. Na Figura 2.7 nos deparamos com retângulos e não tão facilmente com retas.

Figura 2.7 - Ilustração da Lei da forma fechada

- **Lei da boa continuação, destino comum, fato comum ou da boa curva** – Estímulos tendem a ser agrupados juntos de tal maneira a minimizar mudanças ou discontinuidades. As partes que têm “destino comum” são mais facilmente percebidas como unidades do que aquelas que dele carecem, conforme pode ser visto na figura 2.8.

Figura 2.8 - Ilustração da Lei do destino comum

- **Lei da simplicidade (Lei da “boa figura” ou Lei da pregnância)** – Esta é a lei fundamental da psicologia Gestalt. Koffka assim a descreveu: “A organização psicológica é tão *boa* quanto o permitam as condições dadas e o organismo tende a determinar comportamentos proeminentes seja nas percepções, nos movimentos ou nas ações”. Por *boa* organização que procuramos atribuir às coisas, entendemos tudo o que decorre da simetria, regularidade, homogeneidade, coesão, equilíbrio, simplicidade e concisão, apresentadas aos nossos olhos por um todo ou um conjunto. Particularmente, a um exame superficial na Figura 2.9, o círculo e o quadrado apresentam-se como inteiros. Estímulos ambíguos tendem a ser resolvidos em favor da alternativa mais simples. Por exemplo, diferentes projeções de um mesmo cubo são percebidos como bidimensionais ou tridimensionais dependendo de qual interpretação é a mais simples.

Figura 2.9 - Exemplo da lei da boa figura.

- **Lei da simetria** – Quanto mais simétrica uma figura, mais ela parece ser uma figura. Quanto maior o número de simetrias que uma região possui, maior será essa tendência (figura 2.10).

Figura 2.10 – Exemplo da lei da Simetria – Regiões definidas por bordas simétricas tendem a ser percebidas como figuras coerentes

- **Lei do fato comum** – Se um grupo de pontos estão se movendo com uma velocidade uniforme através de um campo de pontos estacionários similares, os pontos móveis são percebidos como um grupo coerente.
- **Lei da área** - Quanto menor a região fechada mais ela tende a ser vista como uma figura. O complemento também assegura que quanto maior a área de uma região, mais ela parece ser o fundo (ver figura 2.11)
- **Lei da convexidade:** Formas convexas tendem a ser vistas como figuras e usualmente sobrepõem simetria.
- **Lei da orientação:** Orientações vertical e horizontal são mais provavelmente vistas como figura do que outras orientações.
- **Lei da Experiência:** Além da experiência biogenética (aquela que o sistema nervoso aprendeu durante a evolução e que tende a constituir formas) a experiência individual ajuda a modelar nossas impressões. A experiência individual é entendida como a consolidação da reação natural do sistema óptico.

Efetivamente, o organismo tende a comportar-se de maneira “proeminente”, não considerando os detalhes e as particularidades, chamando em causa o essencial. Se as formas da Figura 2.9 fossem apresentadas em frações de segundo ou em um ambiente de pouca luminosidade, ou bastante diminuídas em tamanho ou à uma certa distância, elas seriam julgadas inteiras. Em todos esses casos se evidencia com clareza a lei da pregnância.

O aspecto interessante da teoria Gestalt é o postulado de que existe uma tendência intrínseca para organizar em configurações aquilo que é percebido e essa organização é feita de acordo com alguns princípios de tal forma que aquilo que é percebido pode não corresponder exatamente ao estímulo externo. A organização ocorre de acordo com a Lei da Boa Forma segundo a qual as configurações tendem a parecer tão claras, nítidas e estáveis quanto seja possível. Isso pode ocorrer através da simplificação da figura pelo aumento da simetria e regularidade. Os fatores de fechamento, continuidade, inclusão e boa articulação podem atuar de tal maneira que integrem as partes da figura em um “*todo*” integrado. Mesmo quando o estímulo consiste de partes dispersas, haverá uma tendência para associá-los, seja através de proximidade ou através de semelhança. Seria possível supor que as crianças mais jovens que percebem menos claramente que os adultos as características exatas da forma, estariam mais sujeitas aos efeitos da lei de *prägnanz* (Vernon, 1974). Quanto à simplicidade, isso é provavelmente verdade pois a minúcia excessiva é desprezada por elas. Quanto à globalidade de percepção, Elkind et al, (1964) apresentaram às crianças de quatro à nove anos de idade figuras cujas partes constituintes eram formadas por outras figuras representativas, por exemplo cenouras e folhas eram reunidas para formar um avião. As crianças menores viam mais facilmente as partes do que o *todo*. Essa tendência diminuía até os nove anos e os *todos* eram vistos cada vez em proporção maior. Todavia, as partes pareciam mais integradas do que os “*todos*”.

Nenhuma dessas leis são tão simples quanto parecem. O agrupamento, por exemplo, parece estar relacionado com medidas do espaço percebido (em oposição a proximidade medida pela distância na retina) e é influenciado pela experiência anterior.

Apesar das leis de organização se aplicarem bem em alguns exemplos, a operação de algumas dessas leis nem sempre é direta como nos exemplos. Considerando por exemplo a lei da simplicidade, como podemos dizer que uma figura é mais simples do que outra? Infelizmente a lei da simplicidade não provê informação suficiente sobre o que é simplicidade. Uma outra questão sobre a teoria Gestalt seria em como aplicá-la em situações da vida real. A grande dificuldade com a teoria estava relacionada de que embora ela parecia estar intuitivamente correta, não houve nenhuma tentativa em especificar ou quantificar exatamente em uma forma matemática os mecanismos fundamentais ou como eles funcionavam. Um trabalho nesse sentido é mostrado em Zobrist & Thompson, (1975) que abordou esse assunto do ponto de vista matemático, construindo uma medida de distância entre os agrupamentos Gestalt. Outros trabalhos nesse sentido podem ser vistos em (Sakar

& Boyer, 1993). Uma exposição detalhada da teoria Gestalt pode ser encontrada em Werthermeier, (1923), Ellis, (1923), Kohler, (1929), Koffka (1935).

Zusne (1970 a) propôs uma lista de leis para as formas visuais (Loncaric, 1998):

- A forma visual é a propriedade mais importante de uma configuração.
- A forma visual é dinâmica ou resulta de processos dinâmicos nos quais se baseia.
- Todas as formas visuais possuem no mínimo dois aspectos referentes ao fundo e à figura.
- As formas visuais possuem um ou vários centros de gravidade sobre os quais a forma está organizada.
- As formas visuais são transponíveis (com respeito à translação, orientação, tamanho e cor) sem perda de identidade.
- As formas visuais resistem às mudanças. Elas tendem a manter sua estrutura contra forças perturbadoras.
- As formas visuais serão sempre tão boas (regular, simétrica, simples, uniforme, exibindo o mínimo de pressão) quanto as condições (estímulos) permitam.
- As formas podem se fundir para produzir novas formas.
- Uma mudança em uma parte da forma afeta todas as outras partes da forma (Lei da compensação).
- As formas visuais tendem a aparecer e desaparecer como um “todo”.
- As formas visuais deixam um pós – efeito (*aftereffect*) a fim de torná-las fáceis de lembrar (Lei da reprodução)
- O espaço é anisotrópico⁶, possuindo diferentes propriedades em diferentes direções.

2.2.2 - Dependência funcional da figura e fundo

Quanto às características, a figura depende do fundo sobre o qual aparece. O fundo serve como uma estrutura ou moldura em que a figura está enquadrada e por conseguinte determina a figura.

A Figura 2.11 mostra uma figura ambígua extraída de Koffka, (1935). Quando comparadas entre si as partes que constituem a figura e o fundo, verificamos que os fundos

⁶ Anisotropia: Qualidade peculiar a certas substâncias cristalizadas, de reagir diferentemente segundo a direção de propagação de um determinado fenômeno físico, como a propagação da luz.

sempre são mais simples, no sentido de maior uniformidade e menor articulação do que as figuras. Nas figuras ambíguas o mesmo lado pode ser tanto o exterior como o interior, mas quando é um não pode ser outro.

Figura 2.11 – Figura –fundo – (Koffka, 1935)

Um interessante trabalho que aborda a organização figura-ruído e as respostas espaciais e temporais do sistema visual pode ser visto em Weisstein & Wong (1986).

2.3 – Teoria do comportamento de Hebb (1949)

Uma outra teoria proposta para a forma visual foi o trabalho de Hebb, (1949). Hebb apresentou uma teoria neuropsicológica do comportamento em seu livro “A organização do Comportamento”, onde enfatiza o papel das estruturas neuronais no mecanismo da percepção visual. Seu trabalho influenciou as pesquisas na área das Redes Neurais Artificiais.

Em oposição à escola Gestalt, Hebb argumentou que a forma não é percebida como um “todo” e sim consistindo de partes. A organização e as relações entre as partes devem ser aprendidas para um reconhecimento correto. O aspecto da aprendizagem é o ponto central da teoria de Hebb. O movimento dos olhos é o mecanismo principal que integra elementos mais simples de percepção, como linhas e ângulos. Hebb introduziu também a noção de estruturas celulares (*cell assemblies*) que representam neurônios que foram agrupados juntos devido a disparos repetidos e baixa resistência sináptica, que podem se excitar mutuamente. Sua teoria é quantitativa e não computacional, tornando difícil sua implementação.

2.4 – Abordagem Ecológica de Gibson (1966)

Gibson (1966) desenvolveu uma outra teoria para a percepção da forma. O primeiro princípio de sua teoria é que o espaço não é uma entidade geométrica ou abstrata mas sim um espaço visual real caracterizado pelas formas presentes no mesmo. A teoria de Gibson baseia-se na percepção de objetos tridimensionais reais e não suas projeções bidimensionais. O segundo princípio é que um estímulo do mundo real existe atrás de qualquer percepção

visual, simples ou complexa. Esse estímulo apresenta-se na forma de um gradiente, o qual é propriedade da superfície. Exemplos de gradientes físicos são a mudança de tamanho dos elementos de textura (profundidade), grau de convergência de bordas paralelas (perspectiva), matiz (*hue*) e saturação de cores e sobreamento (*shading*). Gibson argumentou que a escola Gestalt ocupou-se com o estudo de projeções bidimensionais de formas provenientes de um mundo tridimensional, e que seu dinamismo não é nada mais que ambigüidade de interpretação de imagens projetadas (Loncaric, 1998).

Egon Brunswik apud Hammond, (1966) estudou a constância de tamanho em ambientes naturais e indicou que através de inúmeras amostras extraídas do ambiente (“amostras ecológicas”), um observador é levado a probabilidades muito grandes, quanto aos muitos aspectos esperados do ambiente, e estas amostras influenciam muito nas indicações ambientais selecionadas para percepção e julgamento. Além disso, através de um *feedback* que verifica a exatidão dos dados sensorados através de suas ações, o observador pode escolher as indicações que mais provavelmente lhe darão informações verídicas quanto à natureza dos objetos e dos ambientes. Isso é muito evidente em julgamentos de distância. Neste caso, o observador pode rapidamente identificar os objetos e julgar suas localizações espaciais a partir de informações imperfeitas e parciais, embora possa não conseguir fazê-lo com exatidão perfeita. No entanto, usualmente existe uma redundância de informação que a reduz a um erro mínimo. As características distais do ambiente precisam ser transmitidas através das características proximais da imagem projetada na retina e não existe uma correspondência de 1:1 entre elas. Os impulsos da retina não podem levar ao cérebro toda a informação necessária para o conhecimento verídico. É essencial que o observador estabeleça relações estáveis ou invariâncias com o ambiente, mas isso só pode ser realizado a partir de base probabilística. As características simples de cor, forma e movimento precisam ser sistematizadas e processadas esquematicamente a fim de dar identificação correta. Pode haver certa perda de exatidão, alguns aspectos podem ser ignorados, deformados ou exagerados, disso decorre a “super constância” de tamanho como resultado da “super compensação” para a distância no ambiente natural.

O método geral de Gibson, (1966) acentua a estreita correspondência entre o que é percebido e os dados sensoriais “invariantes” ligados à algumas propriedades físicas de objetos e seu ambiente espacial. No entanto, embora reduza a um mínimo as funções relacionadas à percepção de lembranças e idéias, acentua a importância da aprendizagem

para o emprego de dados sensoriais, de tal forma que seja mais eficiente para o aperfeiçoamento da discriminação.

Na classificação de Gibson, existem dez tipos diferentes de formas (Loncaric, 1998):

- Forma sólida. (Ver um objeto significa ver uma forma sólida).
- Forma de superfície. (formas com bordas).
- Contorno da forma (perfil) – desenho das bordas de uma forma sólida.
- Forma pictórica – representações (desenhos, fotografias, pinturas)
- Forma plana – desenho das bordas de uma superfície projetada em uma superfície plana.
- Forma perspectiva – desenho em perspectiva de uma forma.
- Forma sem sentido – desenhos que não representam um objeto real.
- Forma geométrica plana – forma geométrica abstrata
- Forma geométrica sólida – Uma parte abstrata de um espaço tridimensional envolto por superfícies imaginárias.
- Forma projetada – projeção de uma forma geométrica plana.

Essas formas podem ser agrupadas em três classes:

- Objetos reais: formas sólidas e superfícies.
- Representações de objetos reais: contorno, representações pictóricas, perspectivas e formas sem sentidos.
- Objetos abstratos (não reais): formas geométricas planas, formas geométricas sólidas e formas projetadas.

2.5 – Processamento Perceptivo

Alguns estudos realizados pela Psicologia Cognitiva a partir da década de 60, foram além dos primeiros estudos da psicologia Gestalt e enfatizaram o processamento perceptivo, onde a percepção é vista como um produto final de operações mentais sobre o estímulo. Psicólogos cognitivos adotaram a idéia de que a mente, como um computador, pode ser vista como um sistema que processa as informações. Essa idéia é análoga à idéia de que neurônios processam informação, entretanto a psicologia cognitiva preocupa-se com eventos nervosos e não com impulsos elétricos. Com essa abordagem, processos cognitivos são quebrados em uma seqüência de passos, conforme mostrado para a memória, na figura 2.12, que é quebrada em Memória de Curta Duração (*Short Term Memory – STM*) e Memória de

Longa Duração (*Long Term Memory - LTM*)⁷. A memória de curta duração é aquela onde as informações ficam disponíveis temporariamente, para que possamos realizar tarefas mais complexas, enquanto que na memória de longa duração as informações ficam armazenadas em caráter permanente.

Figura 2.12 – Exemplo de como um processo cognitivo pode ser quebrado em uma sequência de passos. A memória é quebrada em Memória de Curta Duração (*Short Term Memory - STM*) e Memória de Longa Duração (*Long Term Memory*) (Goldstein, 1989).

A Cognição pode ser definida como o estudo da atividade mental, e se estende a diversas áreas além da percepção, como memória, raciocínio e resolução de problemas. A idéia de que comportamentos tais como atenção, memória e solução de problemas possam ser analisados em uma sequência de eventos mentais cresceu em paralelo com o aumento na popularidade de computadores. Psicólogos cognitivos não apenas analisam comportamentos em diagramas de fluxos como mostrado na figura 2.12, como também escrevem programas para resolver problemas e perceber objetos em cenas simples.

Gregory (1973) apud Goldstein (1989) propôs que a percepção baseia-se no princípio da probabilidade (plausibilidade). De acordo com esse princípio, se um número de objetos possíveis causa um determinado padrão de claro e escuro na retina, iremos perceber

⁷ - As porções do cérebro envolvidas com memória são as áreas de associação dos lobos temporal, occipital, parietal e frontal e partes do sistema límbico, especialmente o hipocampo. Uma teoria para a memória de curta duração é baseada na idéia de “reverberação” dos circuitos neuronais, na qual um impulso circula através de um circuito fechado de neurônios, assim um impulso pode ser lembrado enquanto a reverberação continua. A memória de longa duração pode ser explicada pelo conceito de “facilitação sináptica”, onde a entrada da informação em um circuito neuronal é facilitada devido a passagem prévia de sinais similares.

o objeto mais provável de ocorrer em uma situação em particular. Em alguns casos a escolha de selecionar uma hipótese mais provável é consciente mas testes de hipóteses não ocorrem apenas á nível consciente. Não estamos cientes de muitos processos mentais complexos como por exemplo a percepção de formas e objetos.

As conclusões básicas que as pesquisas em percepção têm chegado é a que a percepção global de um objeto é construída a partir de informações tomadas de partes menores. Para percebermos objetos analisamos seus **componentes ou atributos**, mas ainda não se conhece com certeza quais as unidades básicas de análise usadas pelo sistema visual, conforme pode ser visto nas próximas teorias descritas a seguir.

2.6 – Abordagem Construtivista (1970)

Hochberg (1970) apresentou uma abordagem construtivista para a percepção, onde o observador possui um papel *ativo* na percepção realizando uma série de fixações (sacadas) em diferentes partes de um objeto ou de uma cena. Cada fixação fornece informações (dados) sobre uma pequena área e nossa percepção de todo objeto ou cena ocorre quando combinamos informações de um determinado número de fixações. Os movimentos dos olhos seriam necessários para vermos todos os detalhes de uma cena porque uma simples fixação revelaria detalhes apenas na região próxima da onde se está olhando. Segundo Hochberg, os movimentos dos olhos teriam ainda o objetivo de criar um mapa mental da cena através de um processo de integração de várias informações extraídas. Dessa maneira, a percepção é vista como um processo ativo, no qual a informação é combinada ou processada pela mente do observador. Essa visão da percepção é bastante compatível com a abordagem fisiológica que assume que nossas percepções são construídas a partir de informações contidas nos impulsos nervosos.

2.7 – Teoria Computacional de Marr (1976)

Segundo Marr (1976) o passo básico para analisar uma imagem é a identificação das bordas dos objetos. Marr desenvolveu um paradigma denominado “Esquema Primário” (*Primal Sketch*) para um processamento primário da informação visual (*early processing*). Seu método, através de um conjunto de máscaras é utilizado para medir descontinuidades na primeira e segunda derivadas da imagem original. Essa informação é então processada por procedimentos subsequentes para gerar um esquema primário da cena. O esquema primário

contém localizações das bordas da imagem e é utilizado nos estágios subsequentes de análise de forma. Segundo Marr & Nishihara, (1978) o processamento primário consiste ainda no esquema 2 ½- D é feito uma representação das superfícies visíveis no espaço tridimensional.

Marr & Hildreth (1980) desenvolveram o conceito do *primal sketch* e propuseram um novo filtro detector de bordas baseado no *zero crossings* do Laplaciano da função de distribuição Gaussiana bidimensional. Nessa abordagem, os zeros do Laplaciano indicam pontos de inflexões na borda.

A análise matemática que localiza as bordas dos objetos também permite a extração de primitivas tais como *blobs* (*loops* fechados), segmentos de borda, barras (passagens abertas) e terminações. O esquema primário é o resultado de estágios iniciais de computações, mas nós não o vemos. Antes que a percepção consciente ocorra, o sistema visual deve processar as informações contidas no esquema primário. Primeiro, primitivas que são similares em tamanho e orientação são agrupadas, seguindo o princípio Gestalt e depois são processadas. Os dados utilizados em seu método são as características da imagem da retina, particularmente padrões de áreas escuras e claras na imagem. O computador calcula a existência de objetos no ambiente baseado nesses dados. O sistema proposto por Marr admite que percebemos formas baseado apenas na informação na retina (processamento *bottom-up*) e não em processos *top-down* fundamentado em conhecimento prévio ou em outros tipos de pistas.

2.8 – Teoria Psicológica de Biederman (1987)

Biederman (1987) propôs uma teoria que preocupa-se não em como os atributos de forma e cor são combinados mas em como reconhecemos objetos tridimensionais a partir de seus componentes (primitivas). De acordo com a teoria de Biederman, os seres humanos reconhecem os objetos através da (1) análise (*parsing*) dos objetos em seus elementos constituintes, (2) da determinação da forma que compõem os objetos e (3) da identificação das posições relativas dos componentes, teoria denominada “Reconhecimento-Por-Componentes” (*Recognition- By -Components – RBC*).

Embora a teoria RBC se assemelhe aos métodos de Visão Computacional para a representação de objetos (Marr, 1982) ela é diferente em importantes aspectos. Em Visão

Computacional as formas são tipicamente descritas utilizando-se equações ou estruturas detalhadas de dados, resultando em uma abordagem meticulosa para computar formas (Ballard & Brown, 1982; Marr, 1982). A identificação humana da forma, entretanto é rápida e aproximada. Biederman propôs que o sistema visual humano identifica formas baseado *em propriedades não-acidentais* tais como simetria, paralelismo e linearidade de formas bidimensionais na imagem da retina. Essas propriedades são não-acidentais porque o objeto tridimensional que a produziu também possuiu as mesmas propriedades. As propriedades não-acidentais são usadas pelo sistema visual humano para agrupar estímulos em grupos coerentes formando os componentes de um objeto.

Biederman propôs também que o sistema visual classifica uma forma de entrada em um pequeno número de categorias. Através da combinação de duas ou três propriedades não acidentais Biederman obteve 36 categorias de formas 3-D denominadas *geons*. Diferentes *geons* podem variar em razão de aspecto, e graus de curvatura e paralelismo. A combinação de dois ou mais *geons* podem formar milhões de objetos diferentes, suficientes para representar os objetos rígidos que observamos no mundo real. A figura 2.13 exemplifica alguns *geons* para descrever a forma de partes de objetos.

Figura 2.13 – Exemplos de *geons* (categorias de formas 3-D) proposta por Biederman, (1987) para descrever partes dos objetos. Figura (a) possui uma seção cruzada retangular, enquanto que figuras (b), (c) e (d) possuem seções elípticas. Figuras (a), (b) e (c) possuem eixos de simetria retos, enquanto que os eixos de simetria de (d) é curvo. As seções cruzadas de (c) e (d) aumentam ao longo de seus eixos.

2.9 - Teoria da percepção da Forma de Treisman (1986)

Anne Treisman (1986) argumentou que a habilidade de percebermos um “todo significativo”, aparentemente sem esforços, depende de processos complexos, onde características extraídas automaticamente das cenas são “montadas” em objetos. Treisman propôs que a percepção ocorre em dois ou mais estágios. O estágio “pré-atentivo” (*preattentive*) é rápido e automático e o estímulo visual de entrada é decomposto em propriedades básicas denominadas primitivas (“palavras básicas na linguagem da percepção”). Uma propriedade do estágio pré-atentivo é que ele não requer atenção consciente. Um exemplo de uma tarefa que não necessita de atenção consciente é determinar o “Q” entre os “Os” na figura 2.14 (a). Uma vez extraídas as primitivas, elas são então combinadas no estágio de “atenção focalizada”. Em contraste com o estágio pré-atentivo, esse estágio não é automático e necessita de atenção consciente como ocorre na figura 2.14 (b), onde a busca de um “T” entre “Ls” necessita de busca consciente. A idéia de processamento atencivo foi proposta originalmente por Neisser (1967) que percebeu que alguns aspectos dos estímulos são obtidos automaticamente sem esforço consciente.

Figura 2.14 – É mais fácil determinar “Q” em (a) mas é difícil determinar “T” em (b). De acordo com Treisman (1986) Q é determinado por um processo automático “pré-atentivo” enquanto que o T é determinado por um processo mais lento envolvendo “atenção focalizada” (Goldstein, 1989).

O processo de combinar unidades obtidas no estágio pré-atentivo onde as unidades são analisadas automaticamente e inconscientemente, em um estágio que necessita que o observador focalize a atenção em uma localização particular. Um exemplo de que a atenção *em uma localização* é importante durante o estágio de combinação das características pode ser visto na seguinte experiência realizada por Treisman. Uma tela consistindo de um “X” marrom, vários “Os” vermelhos e vários “Xs” azuis é brevemente amostrada (200 ms) e

seguida de uma tela com pontos padrões (máscara) a fim de eliminar qualquer pós-imagem⁸ (*afterimage*) que possa permanecer quando o estímulo é desligado (figura 2.15). Ao observador era requisitado dizer se o X marrom estava presente e em caso positivo deveriam indicar a localização do mesmo. As pessoas investigadas relataram uma certa dificuldade em determinar a localização do caracter X procurado, apesar de detectarem a presença do mesmo. Segundo Treisman uma razão para essa dificuldade é que as pessoas podem identificar a presença do X marrom apenas baseando-se em uma primitiva (cor). Quando apenas uma primitiva está envolvida, não há necessidade de se focalizar a atenção e a pessoa não presta atenção na localização do objeto. Considere entretanto a tarefa de detectar um único “O” azul entre vários “O”s vermelhos e “Xs” azuis, conforme indica a figura 2.15 (b). Esta tarefa necessita de atenção focalizada em uma determinada localização porque para solucionar este problema o observador tem que localizar um objeto que contém duas propriedades: presença do azul e presença de “O” (combinadas). Assim, quando as pessoas investigadas identificam corretamente a presença de um alvo nesta tarefa, elas também são capazes de indicar sua localização.

Figura 2.15 - Experiências realizadas por Treisman para explicar a busca de uma alvo, X marrom em (a) e O azul em (b). A detecção da presença do X marrom em (a) baseia-se apenas em uma primitiva (cor), enquanto que em (b) o alvo compartilha com os outros elementos duas propriedades presença do azul e presença de “O” . Para combinar as duas primitivas é necessário focalizar a atenção em localizações específicas.

Um outro exemplo de busca visual não instantânea pode ser vista na figura 2.16 a seguir, onde é necessário um tempo maior de busca para se determinar o caracter “R” devido

⁸ Pós-imagem (*afterimage*) é uma ilusão visual causada pelos nossos olhos. Quando olhamos um certo tempo para objetos com alto-contraste, os olhos geram uma imagem repetida do objeto só que nas cores complementares, quando o estímulo não está mais presente.

ao fato do mesmo compartilhar características de linha vertical e linha curva com os outros tipos de elementos (Q e P).

Figura 2.16 – Busca Visual do estímulo R não é instantânea, a menos que ele seja encontrado primeiro ao acaso.

Treisman demonstrou também que algumas bordas entre regiões diferentes se destacavam imediatamente (*pop-out boundaries*). Na figura 2.17 a seguir pode-se perceber uma separação entre as regiões de “Os” e “Vs”, devido a uma diferença de curvatura entre as linhas dos caracteres. *Pop-out boundaries* estão associadas com inclinação das linhas, curvatura e cor (Goldstein, 1989). Medindo-se o tempo levado para as pessoas detectarem as bordas das regiões e encontrar as letras alvos, Treisman identificou as seguintes primitivas: curvatura, inclinação, cor, extremidades das linhas, movimento, áreas fechadas, contraste e brilho. Tais primitivas só serão combinadas no estágio de atenção focal.

Figura 2.17 – Distinção imediata (*pop-out boundary*) entre regiões de “Os” e “Vs”.

Treisman mostrou que essas primitivas existem de maneira independente uma das outras no estágio pré-atentivo utilizando um procedimento denominado técnica da *conjunção ilusória*. Nesse procedimento foi amostrada durante 200 ms uma tela contendo Xs vermelhos, S azuis e T verdes. Observadores relataram terem visto conjunções ilusórias do tipo X verde e S vermelho em um terço das experiências. Esses resultados ocorrem mesmo para estímulos que diferem em forma e tamanho, por exemplos um pequeno círculo

azul e um quadrado verde grande pode produzir uma conjunção ilusória “quadrado azul”. Segundo Treisman isso ocorre devido ao fato de que propriedades do estímulo tais como “presença de vermelho”, curvatura e inclinação são detectadas independentemente no estágio pré-atentivo e como estão disponíveis podem ser incorretamente combinadas quando um estímulo é apresentado brevemente.

Um interessante aspecto relacionado com o agrupamento e a detecção de fronteiras entre regiões distintas pode ser percebido na figura 2.18. Segundo a lei Gestalt de similaridade os caracteres T normais e inclinados deveriam ser agrupados juntos, entretanto isso não acontece. Esse resultado indica que a comparação de um estímulo não é uma maneira segura de determinar quais as propriedades responsáveis pelo agrupamento perceptivo. Na figura 2.18 a propriedade relevante é a orientação.

Figura 2.18 – (a) A orientação é uma força poderosa responsável pelo agrupamento perceptivo (Beck, 1982). Em (b) o elemento-espelho destaca-se em relação a figura normal, enquanto que o elemento-invertido não provoca segregação de textura.

A figura 2.19 mostra um diagrama simplificado do complexo completo para percepção da forma proposto por Treisman (1986).

Figura 2.19 – Diagrama simplificado da sequência de processamento para a teoria de Treisman. O estágio pré-atentivo é quebrado em primitivas e isso é um processo *bottom-up* pois depende apenas das propriedades físicas dos estímulos. Essas primitivas são então combinadas no estágio de atenção focalizada, o qual é afetado tanto pelas informações dos estímulos de entrada provenientes do estágio pré-atentivo quanto pelos processos *top-down* como o conhecimento do observador a respeito do mundo, conhecimento dos nomes e das funções de vários objetos e contexto significativo através do qual o estímulo é visto. O resultado desse processamento é a percepção do objeto como um todo (global).

Outros detalhes sobre as pesquisas de Anne Treisman relacionadas ao processamento pré-atentivo, busca visual e a teoria da Integração das características pode ser visto no Capítulo 3.

A figura 2.20 a seguir mostra o modelo hipotético de Treisman. O sistema visual começa codificando um certo número de propriedades simples e úteis em um conjunto de mapas. No cérebro esses mapas preservam as relações espaciais do mundo visual, entretanto nos estágios subseqüentes essas informações não estão disponíveis e a atenção focalizada seleciona e integra as características em localizações particulares. Nos estágios posteriores, a informação integrada sobre as propriedades e relações em cada arquivo do objeto é comparada com uma rede de reconhecimento. A rede especifica os atributos críticos de árvores, gatos e todos os outros objetos familiares, permitindo acessar seus nomes, seus comportamentos prováveis e seu significado corrente.

Figura 2.20 – Modelo hipotético proposto por Anne Treisman (1986)

2.10 – Rotinas Visuais (1984)

A percepção visual necessita da habilidade de extrair propriedades de forma e relações espaciais entre os objetos e suas partes e essa habilidade é fundamental pois muitos objetos são definidos visualmente pelas propriedades abstratas e relações espaciais entre os objetos. Os seres humanos são capazes de discriminar várias características apresentadas pelo estímulo visual, de reconhecer objetos familiares e de determinar as relações entre os objetos em uma cena visual. O uso da informação visual frequentemente necessita da

extração das propriedades de formato de contorno e regiões e das relações espaciais entre os itens presentes na imagem. A figura 2.21 ilustra um exemplo simples de como o reconhecimento dos objetos dependem não apenas da presença de certas características, mas também de sua disposição espacial.

Figura 2.21 – Desenhos esquemáticos de uma face (a) normalmente disposta e (b) “embaralhada”. A figura (a) é reconhecida imediatamente como face embora as características individuais sejam menos significativas. Em (b) as mesmas partes constituintes estão dispostas em outras posições, dificultando seu reconhecimento (Ullman, 1984).

A análise da forma visual e suas relações espaciais⁹ exercem um importante papel no planejamento de ações e na manipulação de objetos, entretanto não são consideradas no processamento primário (*early processing*), mas possuem grande importância no estágio posterior (*intermediate-level*). É impossível extrair todas as propriedades de forma relevantes (abstração) e todas as relações espaciais em uma imagem de uma maneira *bottom-up*. Diferentes tarefas podem requisitar a extração de diferentes propriedades e suas relações. Essas habilidades exercem um importante papel no reconhecimento de objetos, manipulação guiada visualmente e pensamentos visuais mais abstratos. A manipulação de objetos baseia-se nas informações visuais para responder questões do tipo “A é maior que B?”, “A cabe dentro de B?”. Problemas desse tipo podem ser resolvidos sem necessariamente implicar no reconhecimento do objeto. Ele necessita entretanto, da análise visual da forma e da relação espacial entre suas partes e entre os objetos. Relações espaciais no espaço tridimensional entretanto exercem um importante papel na percepção visual.

⁹ As propriedades de forma tais como orientação, área, etc., referem-se à um único item, enquanto que as relações espaciais tais como *acima*, *dentro*, *maior que*, etc, envolvem dois ou mais itens. Ullman estendeu em seu trabalho o termo propriedades espaciais à propriedades da forma.

A percepção das propriedades espaciais e suas relações, que para o sistema visual humano parece ser realizada de maneira imediata e sem esforços, são complexas do ponto de vista computacional e representam dificuldades fundamentais que afeta todo o processamento da informação visual. A habilidade de estabelecer relações espaciais manifesta-se no sistema visual muito cedo Crianças de 1-15 meses mostraram preferência à formas de rosto em relação a padrões “embaralhados” (Fantz, 1961).

A aparente percepção fácil e imediata das relações espaciais é ilusória e depende de um arranjo de fatores complexos na determinação de determinadas relações espaciais com uma eficiência considerável. Considere por exemplo a tarefa de comparar o comprimento de dois segmentos de retas. Um desenhista poderia realizar essa tarefa medindo o comprimento da primeira linha, armazenar, mediar a segunda linha e comparar as medidas resultantes. Quando duas linhas são apresentadas simultaneamente, é possível comparar seus comprimentos “apenas olhando”.

Ullman (1984) argumentou que a computação das relações espaciais envolvem dois estágios principais. O primeiro estágio é *bottom-up* e é responsável pela criação de certas representações do ambiente visível. Exemplos de tais representações são o esquema primário (*primal sketch*) de Marr (1982) e o esquema 2 ½- D (Marr & Nishihara, 1978). O segundo estágio envolve a aplicação de processos (*top-down*) chamados *rotinas visuais* nas representações construídas no primeiro estágio. As operações elementares fixas do primeiro estágio constituem o conjunto básico de instruções para processos mais complicados. Isso implica em um estrutura espacial e temporal do processo: uma certa operação é aplicada em um certo ponto no campo visual e dependendo dos resultados, uma outra operação é aplicada em um outro ponto. Essas rotinas podem estabelecer propriedades e relações que não podem ser representadas explicitamente nas representações iniciais (abstração¹⁰).

Ullman sugeriu que os requisitos impostos pela análise visual das propriedades de forma e suas relações espaciais são:

- Capacidade de estabelecer propriedades abstratas e relações (*abstração*).

¹⁰Abstrair: considerar isoladamente um ou mais elementos de um todo, separar. **Separar**, considerar isoladamente (coisas que se acham unidas), separar mentalmente para tomar em consideração (uma propriedade que não pode ter existência fora do todo concreto ou intuitivo, em que não aparece), absorver (Dicionário Aurélio – 1988)

Abstrato: que opera com **qualidade** e **relações** e não com a realidade sensível (sensorial). Que expressa uma característica ou qualidade separada do objeto a que está ligada.

Abstract: resumir, desviar a atenção (para), extrato, resumo (Michaelis, 1987).

- Capacidade de estabelecer uma grande variedade de relações e propriedades, incluindo as recentes (*open-endedness*).
- Necessidade de cooperar eficientemente com a complexidade envolvida no cálculo das relações espaciais (*complexidade*).

Um exemplo de abstração de relações espaciais é a determinação das relações dentro/fora. Formalmente, uma propriedade de forma P define um conjunto S de formas que compartilham essa propriedade. A propriedade de fechamento por exemplo, divide o conjunto de todas as curvas dentro das curvas fechadas que compartilham essa propriedade e de maneira complementar, o conjunto das formas abertas. Claramente em muitos casos o conjunto de formas S que compartilham essa propriedade pode ser muito grande e isso torna impossível a comparação da propriedade P com todos os membros de S armazenados na memória. O conjunto de figuras é grande mas dada uma figura isolada é simples testar se ela faz parte de um conjunto, onde uma única figura é *esperada* (parte relevante da imagem) enquanto que relações dentro/fora ou de fechamento podem ser extendidas a um conjunto muito grande. Quando o universo de formas e suas variações é pequeno, o reconhecimento pode ser determinado por simples *template matching*¹¹. Isso significa que um pequeno número de padrões é armazenado e comparado com a figura em questão. Quando o universo é muito grande esse método torna-se impossível. A tarefa de classificação deve ser possível se o conjunto contém certas regularidades¹². Isso significa que o processo de reconhecimento pode ser quebrado em um conjunto de operações de tal maneira que a computação global para estabelecer P é menor que o armazenamento de todas as formas em S .

A percepção de propriedades abstratas pode ser explicada em termos de computações empregadas pelo sistema visual para capturar regularidades que baseiam as diferentes propriedades e relações.

¹¹ *Template Matching* entre figuras planas pode ser definido como a correlação cruzada. Essa definição pode ser extendida a descrições simbólicas dentro do plano. Nesse caso, em cada localização em um plano, um número de símbolos podem ser ativados e um padrão é então um subconjunto de símbolos ativados. Dado um padrão P e um template T , o grau de *match* m é uma função crescente em $P \cap T$ e decresce em $P \cup T - P \cap T$ (quando P “está posicionado” sobre T a fim de maximizar m).

¹² Regularidade: ordem, bom arranjo, estrutura, disposição segundo certas relações, regras.

A capacidade de estabelecer uma grande variedade de relações espaciais e de características implica na necessidade de um mecanismo através dos quais novas combinações possam ser montadas para alcançar novos objetivos computacionais. Pode-se impor tarefas de análise visual do tipo “determine se os elementos verdes e vermelhos estão do mesmo lado da reta vertical”. Segundo Ullman, devido ao fato do sistema visual poder cooperar eficientemente com tais tarefas, sugere a capacidade de criar novos processos a partir de operações elementares.

A complexidade de operações básicas como a ativação da borda implica que diferentes rotinas podem estabelecer diferentes propriedades e relações e usa a operação de ativação da borda tem que compartilhar os mesmos mecanismos e não mecanismos separados. Um caso especial de complexidade ocorre quando necessita-se aplicar a mesma computação em duas localizações espaciais diferentes. A habilidade de realizar uma dada computação em posições espaciais diferentes pode ser obtida através de módulos de processamento independentes em cada posição. Por exemplo, a orientação de um segmento de linha em uma dada localização da cena, parece ser realizada no córtex visual primário independente de outras localizações. Em contraste, as relações mais complexas não podem ser explicadas assumindo um grande número de módulos independentes para cada localização.

Certas restrições devem ser impostas no cálculo das relações espaciais que compartilham operações elementares:

- (1) Propriedades espaciais e relações são estabelecidas pela aplicação de rotinas visuais a um conjunto de representações visuais primárias.
- (2) Rotinas visuais são montadas a partir de um conjunto fixo de operações elementares.
- (3) Novas rotinas podem ser montadas para alcançar novos objetivos de processamento
- (4) Diferentes rotinas podem compartilhar operações elementares.
- (5) Uma rotina pode ser aplicada a diferentes localizações espaciais. Os processos que realizam a mesma rotina em diferentes localizações não são independentes.
- (6) Na aplicação de rotinas visuais, são necessários mecanismos para dar seqüência às operações elementares e para seleccionar as localizações nas quais são aplicadas.

O uso de rotinas visuais para estabelecer propriedades de forma e relações espaciais origina diversos problemas. Um problema computacional geral consiste na escolha de quais propriedades espaciais e relações são importantes para o reconhecimento e manipulação. Considerando-se o algoritmo, o problema consiste em como essas relações

são computadas. Com relação ao mecanismo, o problema consiste em como essas rotinas são implementadas em redes neurais dentro do sistema visual. Os problemas originados pela noção de rotinas visuais são resumidos a seguir:

- (1) *Definição das operações elementares básicas*. Operações básicas do tipo contar interseções da imagem com uma reta em uma direção predeterminada, traçado da borda e ativação da área (segmentação), podem ser utilizadas para estabelecer outras propriedades e relações. O problema consiste na definição de um conjunto de instruções básicas que serviram de base para a composição das rotinas visuais.
- (2) *Integração*. O segundo problema consiste em como operações elementares são integradas em rotinas significativas. O problema possui dois aspectos. Primeiro, os princípios gerais do processo de integração, por exemplo se operações elementares diferentes podem ser aplicadas simultaneamente. Segundo, como rotinas visuais específicas são compostas de operações elementares (alongação, acima, próximo a, dentro/fora, mais alto que).
- (3) *Controle*. As questões nessa categoria referem-se em como as rotinas visuais são selecionadas e controladas, por exemplo, o que controla a execução de diferentes rotinas durante o reconhecimento visual e como a ordem dessa execução é determinada.
- (4) *Compilação*. Como novas rotinas são geradas a fim de alcançar necessidades específicas, e como são armazenadas.

2.10.1 – Rotinas visuais e seus papéis no processamento da informação visual

No esquema sugerido anteriormente, o processamento visual pode ser dividido em dois estágios principais. O primeiro estágio é *bottom-up* e cria alguma representação base para processamentos visuais primários. O segundo estágio consiste na aplicação de rotinas visuais.

As representações visuais primárias (*early processing*) compartilham um número de características fundamentais, não-articuladas¹³, centradas no observador, uniforme e *bottom-*

¹³ Não articuladas: descrições locais que representam propriedades tais como profundidade, orientação, cor, direção de movimento em um ponto. Isso não exclui a possibilidade de que algum processo de agrupamento ocorre para auxiliar a definição de partes distintas e algumas descrições locais não possam acontecer nas representações básicas.

*up*¹⁴. A definição de unidades complicadas maiores e a descrição das relações espaciais entre suas partes, não é determinada nesse nível.

Em contraste com a construção de uma representação base, as propriedades e relações a serem extraídas não são determinadas pelas entradas apenas: para a mesma entrada visual, diferentes aspectos são explicitados em tempos diferentes, dependendo dos objetivos da computação. Ao contrário das representações básicas, a computação das rotinas visuais não são aplicadas de maneira uniforme sobre o campo visual, apenas sobre objetos selecionados. Uma distinção final entre os dois estágios é que o estágio *bottom-up* é fixo, enquanto que as rotinas visuais não, e permitem a extração de novas propriedades e relações.

A criação de representações visuais não para na representações de base. Os resultados obtidos pelas rotinas visuais são retidos temporariamente para uso futuro. Nas representações incrementais, assim denominadas, que utilizam a informação utilizada previamente, representações são também sendo criadas e modificadas no curso de execução de rotinas visuais. Ao contrário das representações de base, as representações incrementais não são guiadas de uma maneira uniforme e não-controlada: a mesma entrada pode originar diferentes representações incrementais, dependendo da rotina aplicada. Rotinas visuais podem então ser descritas como seqüências de operações elementares que são construídas a fim de se alcançar objetivos computacionais específicos. A aplicação seqüencial de operações elementares parecem contrastar com a noção de paralelismo em percepção visual.

O ato de perceber necessita mais do que a existência passiva de um conjunto de representações. Além da criação de uma representação básica, o processo de percepção depende do objetivo computacional corrente. O processo de aplicar a rotina visual apropriada é aparentemente eficiente e suave e contribui para a impressão de que percebemos toda a imagem em um único *glance* (olhar), quando de fato processamos apenas aspectos limitados dela em um dado tempo. Podemos não estar conscientes do

¹⁴O termo processamento *bottom-up* é definido aqui como o processamento determinado totalmente pela imagem. O processamento *top-down* depende de fatores adicionais, tais como o objetivo da computação (tarefa) e não necessariamente do conhecimento específico do objeto).

Psicologicamente, vários mecanismos que parecem estar envolvidos na criação de uma representação base parecem ser *bottom-up*: suas respostas podem ser previstas a partir dos parâmetros do estímulo sozinho. Eles também apresentam forte similaridade em suas resposta em animais acordados, anestesiados e dormindo (Livingstone & Hubel, 1981). O mesmo não ocorre para algumas células do córtex Parietal Posterior conhecido como caminho da *ação*, que não apresentaram respostas em macacos anestesiados.

processamento restrito, mas toda vez que queremos estabelecer fatos novos sobre uma cena, uma resposta é fornecida executando-se uma rotina apropriada. A aplicação de rotinas visuais origina o problema de guiar a atividade perceptiva e de selecionar as rotinas apropriadas em um determinado instante. Com relação a esse problema, muitas teorias usaram a noção de esquemas (Biederman, 1973; Nesser, 1967); Segundo essas teorias, nossa atividade perceptiva pode ser vista como um levantamento de hipóteses a respeito de um objeto específico e então utilizando um conhecimento prévio detalhado sobre esse objeto a fim de confirmar ou rejeitar as hipóteses. Segundo Ullman, a expectativa em algumas condições, pode facilitar significativamente o processo perceptivo, mas seu papel não é indispensável. Percepção pode ocorrer na ausência de expectativa prévia e mesmo quando as expectativas são contraditórias.

O problema de seleção de rotinas apropriadas pode ser abordado, dividindo-se o processo em dois estágios. O primeiro é a aplicação de *rotinas universais* que podem ser utilmente aplicadas sobre a cena a fim de fornecer algumas análises iniciais. Elas podem ser capazes, por exemplo, de isolar algumas partes proeminentes na cena e descrever, talvez não de forma completa, alguns aspectos gerais de forma, movimento, cor, relações espaciais, etc. Essas rotinas universais devem fornecer informações suficientes para permitir uma indexação inicial para uma memória de reconhecimento, que serve para a aplicação de rotinas mais especializadas.

A utilidade potencial das rotinas visuais nas fases iniciais do processo de reconhecimento é sustentado em parte pelo trabalho de Marr & Nishihara (1978) que demonstrou que uma classificação mais grosseira das formas pode ser obtida pelas rotinas universais, sem conhecimento prévio, independente dos objetos vistos, pode fornecer uma categorização inicial poderosa. Similarmente, Richards (1982)¹⁵ apud Ullman (1984) sugeriu que um conjunto pequeno de atributos visuais (tais como direção do movimento, cor, etc) é freqüentemente suficiente para formar uma boa idéia sobre o que é o objeto (por exemplo, uma pessoa andando) embora a especificação da identidade de um objeto específico (ou qual a pessoa) pode ser consideravelmente mais difícil (Milner, 1974). Esses exemplos servem para ilustrar a distinção no reconhecimento visual entre os estágios universal e específicos. No primeiro, rotinas universais podem ser aplicadas para fornecer informações suficiente

¹⁵ Richards, W., (1982) – “How to play twenty questions with nature and win” - (MIT A.I. Laboratory Memo No. 660) – cambridge, MA:MIT.

para acessar uma categoria mais geral. No segundo, rotinas específicas associadas com uma categoria podem então ser aplicadas. Na ausência de expectativas específicas ou conhecimento prévio, rotinas universais são aplicadas primeiro, seguidas pela aplicação seletiva de rotinas específicas. Resultados intermediários obtidos pelas rotinas visuais são resumidos na representação incremental e podem ser utilizados em rotinas subsequentes.

O papel geral das rotinas visuais no processamento global da informação visual, é intermediar a representação visual e os componentes de alto nível do sistema, tais como memória de reconhecimento. O reconhecimento de objetos familiares, por exemplo, freqüentemente necessita o uso do conhecimento específico sobre os objetos. O que nós sabemos sobre “elefantes” ou “telefones” pode entrar no processo de reconhecimento desses objetos. Em contraste, a extração das relações espaciais, enquanto importantes para o reconhecimento do objeto, é independente do conhecimento específico do objeto. Tal conhecimento pode determinar a rotina a ser aplicada. Um conhecimento completo de problemas tais como o reconhecimento e manipulação visualmente guiada de objetos, requer o estudo adicional de componentes de mais alto nível, como eles determinam a aplicação das rotinas visuais e como são afetados pelos resultados das rotinas visuais aplicadas.

A noção de processamento da informação visual em paralelo, não tem um significado único e bem definido. Existem pelo menos três tipos de paralelismo que podem ser distinguidos nesse processamento: espacial, funcional e temporal. O paralelismo espacial significa que as mesmas operações ou operações semelhantes podem ser aplicadas simultaneamente em diferentes localizações espaciais. As operações realizadas pela retina e pelo córtex visual primário, por exemplo, caem dentro dessa categoria. O paralelismo funcional significa que diferentes computações podem ser aplicadas simultaneamente a mesma localização. Zeki (1978) sugeriu que diferentes áreas do córtex extra-estriado processa diferentes aspectos da entrada tais como cor, movimento e disparidade estereoscópica, na mesma localização simultaneamente, determinando assim paralelismo funcional¹⁶. Paralelismo temporal é a aplicação simultânea de diferentes estágios de processamentos à diferentes entradas (*pipelining*). A aplicação de rotinas visuais é compatível com os três tipos de paralelismo, entretanto o uso de paralelismo espacial é mais restrito nas rotinas visuais do que na construção das representações de base¹⁷. Existem certas

¹⁶ Embora a noção de uma área cortical para cada função seja muito simplista, os dados psicológicos sustentam em geral a noção de paralelismo funcional.

relações que são inerentemente seqüenciais, no sentido que é impossível ou altamente caro aplicar um paralelismo espacial extensivo em sua computação. Nesse caso o comportamento seqüencial é essencial e é imposto pela natureza da tarefa. No processamento seqüencial essencial, existe a necessidade de um mecanismo para controlar a seqüência apropriada de computação. Em processamentos seqüenciais não essenciais a ordem não possui importância.

Em várias teorias da Atenção Visual a ordem seqüencial na percepção é não essencial e origina-se primariamente da capacidade limitada do sistema (Holtzman & Gazzaniga, 1982). Nesse ponto de vista, apenas uma região limitada da cena visual (1 grau – Eriksen & Hoffman, 1972) é processada de uma vez devido a capacidade limitada do sistema. Embora operações elementares sejam seqüenciais, algumas delas como a ativação de borda, aplicam paralelismo espacial e não são confinadas a uma região limitada. Claramente quando o processamento limitado a uma pequena região do espaço, a tarefa torna-se consideravelmente mais difícil. O paralelismo espacial pode exercer um importante papel nas rotinas visuais.

2.10.2 – Operações elementares

Um requisito fundamental na execução de rotinas visuais é a capacidade de controlar a localização que certas operações serão realizadas. Por exemplo, a operação de ativação de área (segmentação) será de pouco uso se a ativação começar simultaneamente em qualquer lugar. Para ser útil ela deve começar em regiões selecionadas, ou ao longo de um contorno selecionado. Mais genericamente, a aplicação de rotinas visuais necessitem de um mecanismo de direcionamento que permita a aplicação de uma mesma operação em diferentes localizações espaciais.

O direcionamento do foco de processamento, ou seja, da área onde serão aplicadas as rotinas visuais, pode ser determinado em parte movendo-se os olhos (Noton & Stark, 1971)¹⁸, entretanto muitas relações por exemplo a relação dentro/fora, pode ser estabelecida

¹⁷A noção geral de um estágio extensivamente paralelizado, seguido de um estágio seqüencial, está de acordo com várias teorias de percepção visual (Estes, 1972; Neisser, 1967; Shiffrin et al. 1976).

¹⁸ Segundo Noton & Stark, (1971) existem algumas indicações da habilidade de se começar a explorar uma figura, iniciando pelo seu centro.

sem o movimento dos olhos. A capacidade de trocar o foco de processamento internamente é necessária.

Problemas relacionados com as possíveis trocas de operações internas, tem sido estudados pela Psicofísica e Psicologia de maneira empírica. Esses estudos diversos não forneceram uma descrição completa dessas trocas de operações e seus usos na análise da informação visual. Eles fornecem entretanto fortes bases para a noção de que a troca do foco de processamento exerce um importante papel no processamento da informação visual, começando nos primeiros estágios de processamento.

Estudos psicológicos sugeriram que o foco de processamento visual pode ser direcionado tanto voluntariamente ou através da manipulação dos objetos em diferentes localizações no campo visual. Muitos dos estudos psicológicos utilizam o conceito de atenção seletiva. Esse conceito possuiu uma variedade de significados e conotações (Estes, 1972). A noção de atenção visual seletiva sempre implica que o processamento da informação está restrita a uma pequena região do espaço, a fim de não sobrecarregar o sistema com informação excessiva como nos processos de coloração (segmentação). Em contraste, a noção da troca de processamento discutida acima origina-se da necessidade de processos estruturados espacialmente e isso não implica necessariamente em noções de capacidade geral ou proteção contra “sobrecarga” de informação excessiva e sim em uma coordenação do processo. Como resultado, o processo é espacialmente estruturado, não de uma maneira simples como o modelo focal da atenção seletiva (*spotlight model*).

Os mecanismos relacionados com a troca da atenção (*shifting*) têm sido estudados em macacos em um grande número de áreas visuais como o Colículo Superior, Lóbulo Parietal Posterior (área 7), campos frontais dos olhos, áreas V1, V2, V4, Temporal médio, Temporal médio Superior e o Lóbulo Temporal Inferior. Nas camadas superficiais do Colículo Superior dos macacos, muitas células apresentaram aumento em suas respostas quando o macaco usou o estímulo como alvo e realizou uma sacada (movimento dos olhos). Esse fenômeno foi denominado de direcionamento da atenção visual (neuronal) uma vez que modifica a entrada visual e a intensifica em localizações selecionadas.

Na determinação da relevância dos estudos fisiológicos é importante a distinção entre três tipos diferentes de interações entre as respostas fisiológicas e a tarefa visual realizada pelo animal: dependência da tarefa, dependência da localização da tarefa e

dependência da localização. Algumas unidades na área TE (Lóbulo Temporal) são claramente dependentes da tarefa, enquanto que unidades na área V1 por exemplo parecem ser independentes da tarefa. Respostas dependentes da tarefa sugerem, que unidades não pertencem a geração *bottom-up* das representações visuais primárias e que podem participar das rotinas visuais. A dependência da tarefa não implica necessariamente entretanto, na existência de operações de troca. Ao contrário da metáfora do ponto focal da atenção visual, pode haver diferentes respostas quando diferentes tarefas são realizadas nas mesmas localizações.

As considerações computacionais sugerem o uso de trocas internas do processo de foco. Essas operações são chamadas de indexação (*indexing*) e pode ser descrita como a troca de atenção para uma localização especial selecionada (saliente). Essas localizações são determinadas em paralelo sobre as representações bases e podem servir como pontos para aplicação de rotinas visuais. Em uma cena nova, é possível direcionar o foco do processamento imediatamente para um item saliente indexado e começar o processamento naquela localização. A operação de *indexing* pode ser dividida em três estágios sucessivos. Primeiro, propriedades usadas na indexação, tais como movimento, orientação e cor, devem ser comparadas sobre as transformações bases. Segundo, um processo de seleção é necessário para definir localizações que são diferentes de sua vizinhança. O terceiro estágio consiste na troca do processamento para as localizações indexadas. Utilizando técnicas que medem tempo de reação de respostas em tarefas de busca visual, Treisman (1977) e Treisman & Gelade (1980) demonstraram que cor e características simples de forma podem ser indexáveis. Outros candidatos incluem curvatura, orientação e número de extremidades.

A noção de um conjunto limitado de propriedades que podem ser processadas pré-ativamente concorda com os estudos de Julezs (1981). Segundo Julezs um conjunto limitado de características denominados “*textons*” podem mediar discriminação imediata de textura. Esses *textons* incluem cor, *blobs* (regiões) alongadas ou de tamanho específicos, orientações, razão de aspecto e terminações das regiões alongadas.

Esses estudos psicológicos concordam com a evidência fisiológica. Propriedades tais como movimento, orientação e cor são detectadas em paralelo por unidades que cobrem todo o campo visual. A figura emergente é, em conclusão, um pequeno número de propriedades computadas em paralelo sobre as representações bases antes da aplicação das rotinas visuais e então projetadas em seqüências de mapas retinotópicos. Várias dessas

propriedades são conhecidas mas a lista completa precisa ainda ser estabelecida. Os resultados são então utilizados em um grande número de tarefas visuais, incluindo discriminação de textura, correspondência entre movimentos, estéreo e indexação.

Seguindo-se a computação inicial das propriedades elementares, o próximo estágio na operação de indexação necessita de comparações entre propriedades computadas em diferentes localizações, a fim de se determinar as localizações indexáveis (salientes). Evidências psicológicas sugeriram que apenas comparações simples são usadas nesse estágio. Vários estudos realizados por Treisman examinaram o problema em como várias características medidas em uma dada localização, podem ser combinadas antes da operação de indexação. Um número de propriedades elementares são computadas em paralelo, comparações locais são realizadas em cada lugar e as diferenças de sinais resultantes são combinadas a fim de produzir um sinal em cada localização. O foco de processamento muda para a localização mais forte.

Além das operações de *shifting* e *indexing*, Ullman definiu ainda as operações de:

- *Ativação delimitada (bounded activation)* ou coloração (*coloring*). Essa operação consiste no espalhamento da atividade sobre a superfície na base de representação, emanando de uma dada localização ou contorno e parando em bordas descontínuas. Essa operação não é simples, uma vez que devem lidar com a existência de contornos externos e bordas fragmentadas. Uma discussão dos mecanismos envolvidos nessa operação sugere que podem existir redes especializadas dentro do sistema visual, para a execução e controle da aplicação das rotinas visuais. Os resultados da operação de coloração devem ser retidos no estágio incremental para uso por rotinas adicionais. A coloração fornece uma maneira de definir unidades maiores na representação não-articulada de base.
- *Traçado do contorno (boundary tracing)*. Essa operação envolve tanto o traçado de contornos simples, ou a ativação simultânea de um número de contornos. Essa operação deve ser capaz de lidar com traçados incompletos de bordas, traçados sobre interseções e pontos ramificados e traçado de contorno em diferentes escalas.
- *Marcação (Marking)*. Na aplicação de rotinas visuais, o processamento pode ser deslocado através de diversas localizações nas representações base. A operação de marcar uma localização significa que a localização é lembrada e o processamento pode ser retornado a ela se necessário. Tais operações são úteis na integração da informação no processamento de diferentes partes de uma cena completa. Uma outra operação que se beneficia é a contagem visual que consiste em se determinar da maneira mais rápida possível o número de itens na

imagem. Quando o número de itens é quatro ou menos, a percepção é tão imediata e direta, relacionado com mecanismos Gestalt (Atkinson et al., 1969). A contagem pode também ser realizada por simples rotinas visuais tais como *shifting* e *marking*.

Essas operações sugerem que mesmo aparentemente simples e imediata, a percepção das relações espaciais refere-se a um conjunto complexo de processos envolvidos na seleção, montagem e execução das rotinas visuais. O reconhecimento de um objeto específico pode usar rotinas pré-montadas para inspecionar características relevantes e relações entre eles. Tais rotinas podem também ser geradas eficientemente por um mecanismo na resposta de objetivos específicos. Pode-se por exemplo armazenar rotinas na memória na forma de esqueleto apenas. Os mecanismos montados irão então detalhar e gerar rotinas intermediárias quando necessário. Em tal esquema, a atividade perceptiva durante o reconhecimento será guiada por um conjunto de objetivos pré-armazenados que o processo de montagem irá expandir em rotinas visuais detalhadas.

O traçado do contorno pode proceder ao longo do esqueleto principal de um contorno sem traçar seus componentes individuais. Um exemplo é ilustrado na figura 2.22 a seguir. Exemplos semelhantes têm sido usados para argumentar que a percepção “global” ou “holística” precede a extração de características locais. De acordo com o esquema das rotinas visuais, linhas são extraídas no processo primário e representado nas representações bases. Os elementos constituintes não são reconhecidos uma vez que o reconhecimento envolve a aplicação de rotinas visuais. O fenômeno denominado “*A floresta antes das árvores*” (Johnston & McLelland, 1973; Navon, 1977; Pomerantz et al., 1977) é o resultado da aplicação de rotinas apropriadas que possam traçar e analisar agregados sem analisar seus componentes individuais, conduzindo o reconhecimento da figura global antes do reconhecimento de suas partes. A habilidade de extrair características da disposição de tokens está relacionada com o papel do agrupamento nos processos primários e participa do processo da formação da representação de base.

Figura 2.22 – A figura global pode ser traçada sem reconhecer todos os seus componentes individuais (Ullman, 1984).

2.11 – Teorias baseadas em Esquemas

A teoria de Esquemas tem sido utilizada por muitos cientistas para explicar funções cognitivas e perceptivas (Head & Holmes, 1911; Piaget, 1952; Arbib, 1989). Muitos pesquisadores acreditam que o entendimento do comportamento complexo e funções cognitivas necessita do estudo de seus circuitos neuronais (Arbib, 1989, Grossberg, 1980; Kossilyn, 1987); Estudando as funções cognitivas apenas no nível funcional, relevando a estrutura e fisiologia dos córtices e neurônios pode levar a uma direção errada.

Segundo Rumelhart (1980) um esquema é um pacote que contém conhecimento sobre o mundo e informações sobre como usar o conhecimento. Diferentes esquemas carregam conhecimento sobre a descrição de objetos, situações, eventos, sequência de eventos, ações e ações subsequentes. Um esquema pode envolver vários subesquemas. Através do uso de esquemas, sistemas inteligentes (biológicos ou artificiais) podem determinar por exemplo se um objeto está presente na imagem. Por exemplo um esquema para o objeto faca pode conter duas variáveis comprimento e localização. Arbib (1989) propôs dois tipos de esquema: motor e perceptivo que segundo ele estão sempre interagindo: o sistema perceptivo obtém novas informações que serão utilizadas para refinamento e modificação dos planos de ação. O esquema motor planeja e o ciclo percepção-ação continua, assim a troca de atenção é vista como uma ação guiada pela percepção. Arbib propôs também que podem ocorrer interações entre os esquemas em termos de cooperação e competição.

Evidências psicológicas e neurobiológicas sobre o comportamento pode direcionar teorias cognitivas de forma. Recentes interesses em abordagens conexionistas e de Redes Neurais em modelar o comportamento procura remediar os inconvenientes de abordagens puramente funcionais (*top-down*), uma abordagem puramente *bottom-up* representa um nível muito baixo para modelar funções complexas no cérebro. É muito difícil considerar dados psicológicos complexos utilizando por exemplo uma simples rede *feedforward* de três camadas. Em contraste é possível lidar com restrições de alto nível, tais como a incapacidade de processar uma cena inteira de uma vez e a necessidade de direcionar a atenção em diversas partes da cena (Didday & Arbib, 1975; Feldman, 1985).

Alguns pesquisadores propuseram modelos intermediários fornecendo uma ponte de ligação entre comportamento e *hardware* neural, como em Arbib (1989). Em particular Arbib (1993) propôs o uso de esquemas como unidades funcionais intermediárias de funções cognitivas

2.12 – Percepção canônica dos objetos

Conforme mostrado nas sessões anteriores, a tarefa básica da percepção visual é extrair informações sobre as propriedades dos objetos no ambiente em termos de padrões de luz. Existem muitos aspectos que dificultam essa tarefa, como por exemplo o número infinito de diferentes imagens bidimensionais de um objeto quando visto em diferentes posições. Apenas uma parte da imagem tridimensional do objeto é visível para uma dada perspectiva e essa porção varia de uma perspectiva para outra.

Palmer et al. (1981) consideram que a estrutura de conhecimento para o objeto está acessível ao máximo de uma perspectiva canônica (fundamental) ou privilegiada. Estudos em cognição e percepção sugerem esse fato, conforme revisto a seguir.

Dentro do contexto de processamento de informação, o reconhecimento do objeto é visto ocorrer quando o estímulo “acessa” ou faz contato com a representação interna apropriada do objeto. Acredita-se que essa estrutura de representação interna contém informação sobre todo o objeto e que a informação do estímulo é mapeada dentro dessa estrutura de conhecimento como um subconjunto (Misnky, 1975). A partir dessas suposições e também devido ao fato de que diferentes vistas revelam diferentes informações

sobre o objeto, percebeu-se que algumas vistas possuem melhor acesso às representações internas do correspondente objeto do que outras. A vista que contém a informação mais significativa de máxima saliência deve ser reconhecível ao máximo.

Pesquisas em categorização sugerem que categorias possuem exemplares “protótipos” com *status* privilegiado. Tais membros possuem mais atributos em comum com outros exemplares e contém mais informação da categoria como um todo (Mervis & Rosch, 1981). Se considerarmos o conceito de um objeto individual perceptível como sendo uma categoria e suas vistas como exemplos dessas categorias, deve haver algo nessas vistas que seja mais típico de tal maneira que seja similar às outras vistas do mesmo objeto mas que contenha a informação mais importante e de maior saliência do objeto, ou seja uma perspectiva canônica.

Meeleau-Ponty (1962) argumentou que um objeto estável com propriedades objetivas emerge apenas quando o campo espacial tenha se organizado, de maneira que o observador tenha uma atenção máxima (“*maximum grip*”) em um objeto.

Palmer et al. (1981) realizaram uma série de experiências a fim de estabelecer uma medida de importância (*canonicalness*) entre várias perspectivas e a informação subjetiva do objeto que elas contêm e concluíram que a familiaridade exerce um papel fundamental em se determinar a importância de vários aspectos dos objetos. A importância dos objetos também está fortemente associada com os aspectos de visibilidade do objeto. O mesmo objeto visto na mesma direção, porém em distâncias diferentes produzem diferentes performances de reconhecimento.

Em Hayward (1998) pode ser vista uma revisão de várias teorias de reconhecimento de objeto que consideram mudanças de perspectiva. Nesse trabalho Hayward investiga também o efeito de reconhecimento dos objetos trabalhando com imagens em níveis de cinza e silhuetas.

3 - EXTENDENDO A PSICOLOGIA GESTALT

A explicação Gestalt para a percepção de forma é partes se tornam “coladas” para criar uma forma coesa, global, se as mesmas atendem aos critérios de similaridade,

proximidade, boa continuação e simetria. Essas partes serão organizadas se puderem se ajustar a “boa figura”, mesmo que isso necessite alguma distorção (não veracidade) (Pomerantz & Kubovy, 1986). Uma vez tão organizada, a percepção global ocorre primeiro e partes só serão atendidas como partes apenas através de um ato secundário de investigação (*scrutiny*) que requer esforço e prática. Segundo Köhler (1920) esse processo global é orquestrado através de operações automáticas de campos elétricos no cérebro. Esse processo trabalha em paralelo sobre o campo perceptivo, estando em grande parte “construído” na infância e portanto não educado através da experiência. O processo pré-atentivo de Neisser (1967) realiza essa mesma função.

As teorias apresentadas nas seções anteriores descrevem a percepção como um processo que extrai unidades básicas denominadas primitivas, e essa conclusão entra em conflito com a idéia Gestalt de que percebemos a globalidade. A psicologia Gestalt baseia-se em evidência introspectiva. Os psicólogos examinaram suas próprias experiências e concluíram que estamos mais conscientes do objeto global do que de suas partes e que os mecanismos básicos de percepção da forma devem ser *holísticos*. O fato de não estarmos conscientes das unidades básicas não significa que elas não exercem um importante papel na percepção. O processamento inicial em termos de unidades básicas não estão acessíveis à consciência, mas o processamento global já está acessível. No nível fisiológico, a percepção é causada pelo processamento neural que também está escondido da consciência. Dessa maneira, se pensarmos que a percepção começa com as primitivas, os princípios Gestalt são parte do processo que nos conduz a percepção das formas.

A evidência da forma e de suas leis evidencia-se em todos os setores. Por exemplo, no campo da acústica, uma melodia não pode ser interpretada como uma *soma* de tons, pois de fato ela conserva a sua *forma* se a transpomos para uma tonalidade mais alta ou mais baixa. Podemos compreender quem nos fala em meio a um grande alvoroço isolando um som, uma forma entre muitíssimos sons. Podemos constituir também em unidade aquilo que se refere a uma frase, a um discurso, a um problema. Em todos os casos, evidencia-se que o estímulo sensorial local, isolado não tem impressões correspondentes e unívocas, mas que as impressões são estruturadas e organizadas através de excitações diversas, oriundas de vários setores sensoriais.

Alguns psicólogos argumentaram que a forma se fundamenta na percepção de relações entre as várias partes ou estímulos recebidos: um motivo musical por exemplo,

seria compreendido em consequência da percepção das relações entre os vários sons. Os adeptos da Gestalt respondem que uma melodia é percebida com imediata simplicidade: a análise de todas as relações existentes entre os tons é inadmissível devido ao tempo que deveríamos procurá-los e compreendê-los num todo integrado. Outros opositores sustentaram que a forma seria um “fenômeno suplementar”. Os psicólogos da Gestalt respondem que a forma não é uma estrutura aditivamente obtida a partir de elementos primários e sim apresenta-se de imediato e não precisa desenvolver nenhuma atividade para percebê-la, atividade essa que seria necessária no caso da produção suplementar.

Afora a sua “totalidade” ou “imediatismo”, as formas são mais ou menos delineadas, coerentes, estruturadas e organizadas. O círculo é por excelência a forma “forte”, perfeitamente coesa e apresenta-se com um grande imediatismo à nossa percepção do que outras formas. Numa forma, algumas partes podem “valer” muito mais do que outras e podem dar sentido à própria forma de maneira indispensável: por exemplo se no desenho de um rosto que sorri suprimimos a boca a expressão desaparece completamente, ao passo que a expressão não será prejudicada se eliminamos um olho.

Devemos sublinhar que podemos *resistir à pressão* das formas e reagrupá-las de maneira diferente da que instintivamente faríamos. Para Köhler, (1929) forma é “um todo delimitado”. Wertheimer, (1945) especifica que um “conjunto é a soma de suas partes ou componentes” somente quando ele puder ser construído com tais partes, sem que em seguida à composição, cada uma das tais partes se modifique. Pelo contrário, as formas são estruturas integrais, cujo comportamento é determinado pela estrutura intrínseca do conjunto. Segundo Koffka, (1937) “a organização é o processo que leva à forma. Esta determinação não seria suficiente para definir a forma, se não lhe incorporássemos a organização, ou seja, precisamente aquilo que é expresso pela lei da pregnância. Essa organização constitui o oposto da distribuição casual”.

Somos conduzidos, por uma certa forma de economia dos recursos mentais, pelas leis da forma, inerentes à organização dos nossos órgãos perceptivos a reduzir a um único objeto, que se torna a *síntese* daquelas impressões. Concentramos assim uma quantidade e variedade de estímulos provenientes de um ou mais órgãos numa única estrutura e conceito, simplificando assim a complexidade do mundo exterior e entre outras coisas, reduzindo o “espaço de armazenamento” na memória. Observamos que se isso é possível, a transformação de um conjunto de percepções em uma figura completa não deve ser muito

diferente da transformação de muitas figuras em uma outra superior, e assim por diante, até as sínteses mais elevadas. Consequentemente, um dos méritos mais relevantes da Psicologia da Forma seja exatamente o de ter identificado experimentalmente os primeiros mecanismos fenomenológicos da origem e da constituição dos conceitos. Com relação às informações precedentes, a descoberta da forma permite o início de uma viagem com nova clareza rumo ao pensamento mais sofisticado tendo à disposição um mecanismo altamente qualificado.

A Psicologia da Forma sustenta a necessidade de uma energia, por exemplo o impulso dado por uma necessidade, por um interesse. Sem essa energia, a contigüidade espaço-temporal não seria suficiente para estabelecer o contato entre elementos psíquicos. Por experimentos se pode demonstrar que apresentando mesmo inúmeras vezes à uma pessoa um determinado material para ser retido na memória, esta não é capaz de fazê-lo a não ser que tenha a precisa intenção de aprender, a menos que tenha outra motivação. Em tais casos, o material é facilmente “organizado” ou “transformado estruturalmente”, inclusive com a contigüidade espaço-temporal. Concluindo, existem em nós vestígios mnemônicos¹⁹, mas eles reproduzem um processo estrutural, organizado e não imagens. É a impressão vivida, estruturada que nos permite reviver imagens²⁰.

3.1 – Gestalt e Atenção

Como se vê os processos volitivos (que dependem da vontade) adquirem um enfoque especial na psicologia da Gestalt. A psicologia tradicional partia dos estímulos, dos reflexos, derivando os atos da vontade através das facilidades, reforços e inibições de várias espécies, sem deixar lugar para os processos espontâneos, alimentados por energias centrais. De acordo com a Teoria Gestalt, o campo perceptivo se auto-organiza e se estrutura automaticamente, mesmo que o sujeito que vive a percepção esteja desinteressado e assume uma atitude completamente passiva. O estímulo existe mas não gera ação se não existe uma

¹⁹ Mnemônico: relativo à memória, associação daquilo que deve ser memorizado com dados conhecidos ou vividos.

²⁰ A maneira que guardamos as informações nos neurônios pode ser considerada uma Gestalt, uma vez que é necessária uma estrutura de ligações sinápticas (facilitações) que permitem o acesso às informações. Ver a discussão do “*mind-body problem*”, apresentada no apêndice 1, que aborda como os processos físicos tais como impulsos nervosos (provenientes de 10 bilhões de células nervosas aproximadamente) interagem para produzir processos mentais tais como experiência perceptiva em uma consciência única (integrada). Segundo Goldstein, (1989): “*Perception is shaped by neural processing*”.

tensão baseada na necessidade interior do sujeito que de certo modo determinará a utilização do próprio estímulo. Por exemplo, a estrutura de uma reação é diferente conforme se trate de um cão faminto ou um cão sem fome. Outros exemplos seriam a caixa de correio que nos atrai quando levamos uma carta no bolso, mas é um objeto indiferente quando a carta já foi despachada ou a campainha do telefone que foi escolhida como um sinal por sua intensidade, repetição e caráter inesperado da campainha. Premidos pela necessidade, as representações e os estímulos impelem o indivíduo a agir em direção à satisfação das necessidades, sejam físicas ou motoras, ou sejam psíquicas quando os estímulos correspondem às necessidades mentais. Todas essas características foram enumeradas como “*condições de atenção*”, entre muitas outras qualidades. Certas qualidades como gosto amargo, cheiro de almíscar e cor amarela exercem um efeito particularmente forte sobre a atenção.

Os estudos das condições da atenção foram considerados desde o início do século XX. Müller (1896) definiu *impressividade* como o poder (força) que as impressões sensoriais atraem (impressionam) a nossa atenção. Titchener, (1910) apud Koffka, (1937) tratou a atenção como uma mera propriedade, atributo ou dimensão dos objetos no campo e privou-a de seu principal carácter que segundo Koffka, (1937) são as relações “Ego-objeto”. Definindo a atenção como uma relação Ego-objeto, pode se fazer jus tanto à atenção voluntária quanto à involuntária. Na primeira a força é oriunda do Ego; na segunda ela provém primordialmente do objeto. Stout, (1909) definiu a atenção como “a direção do pensamento, para este ou para aquele objeto especial, em preferência à outros objetos”.

A atenção não pode ser provocada diretamente pelos estímulos e sim pelos objetos que por sua vez devem sua existência aos estímulos. Segundo Wertheimer, (1925) problema da significação está intimamente vinculado ao problema de relacionamento do todo com suas partes.

As leis válidas para a percepção são válidas para o pensamento. Quando examinamos um conjunto de idéias, deveremos sofrer os efeitos das leis específicas da forma, além dos efeitos da lei da experiência. Na realidade, notamos que duas idéias nos apareceram com maior evidência do que outras por terem maior afinidade entre si (lei da vizinhança). Caso nos apresente um grande número de idéias, ao reagrupá-las teremos a tendência de perceber aquelas que oferecem similitudes entre si (lei da semelhança); as idéias bem formadas, completas e fechadas nos apareceram mais evidentes do que as abertas

e não delimitadas (lei da forma fechada); as idéias que tem “objetivo comum” nos aparecerão interligadas, contrariamente àquelas que têm objetivo diverso (lei do objetivo comum). As idéias que têm dinamismo próprio serão bem mais percebidas do que as estáticas e as idéias com dinamismo semelhante serão agrupadas como uma unidade em relação às outras de dinamismo diferente (lei do movimento comum). Enfim, as idéias até certo ponto incompletas nos aparecerão completas e serão memorizadas, utilizadas e vividas como unidades perfeitamente estruturadas (lei da pregnância).

Naturalmente, as idéias se manifestarão com maior ou menor evidência de acordo com sua força, conforme a maior ou menor afinidade de uma idéia com outras, segundo o dinamismo, a perfeição, a coerência, coesão, etc. que apresentam.

O agrupamento de partes em formas globais (“*configuração*”) pode ser demonstrado avaliando-se a performance em tarefas que envolvem atenção seletiva e dividida (Pomerantz, 1986). Um conjunto de partes pode estar configurado (agrupado em uma forma global) se respostas que são logicamente contingentes (dependentes) de todas as partes puder ser feita (mais) rapidamente e precisamente do que aqueles contingentes em apenas uma (ou algumas) partes. Quando partes estão configuradas em um todo, existe envolvida usualmente alguma característica emergente que é dependente da disposição das partes, como simetria e fechamento. . Uma característica emergente é definida como uma propriedade que é no mínimo mais saliente perceptivamente do que outras partes, das quais emerge. Existe pouca evidência para a existência de qualquer cola perceptiva que une partes em um todo. Frequentemente, características emergentes são as primeiras características de forma que captam a atenção do observador. As vezes a atenção das características emergentes faz parecer que as partes estão inacessíveis ou “coladas” juntas. Coren & Girgus (1980) estudou o agrupamento perceptivo e concluiu que “itens não precisam parecer estar mais próximos para serem agrupados, eles simplesmente parecem pertencer uns aos outros”.

Noções apresentadas por Sheoard (1964), Lockhead (1966), Garner (1970) e Kahneman (1973) clarearam que um grupo perceptivo ou unidade refere-se a uma unidade de atenção ou processamento de informação. Se duas unidades “pertencem” ao mesmo grupo, isso significa que observadores são capazes de atender e responder àquele grupo, como um grupo.

Pretende-se avaliar agora duas afirmativas. Tarefas que requisitam dividir a atenção para várias partes de um grupo não devem ser mais difíceis (ou até mais fáceis) do que tarefas que permitam a atenção a uma parte, porque da perspectiva do observador o grupo é uma única entidade e não uma mistura de partes. Uma segunda afirmação é que a atenção seletiva em partes dentro de um grupo deve ser dificultada ou mesmo impossível. Essa suposição é equivalente à noção de que partes não estão disponíveis perceptivamente, ou que a floresta esconde suas árvores. Os elementos são agrupados juntos por uma “cola” que evita a “bigamia” perceptiva na qual um elemento pode se casar com mais de um grupo simultaneamente. As afirmativas acima foram sustentadas pelo experimento de Pomerantz & Schwaitzberg (1975) que realizaram experiências de agrupamento com os padrões formados por parênteses: “((“ , “)(“ , “()”, e “))”. Variando-se a distância entre os padrões (proximidade) e variando as tarefas (atenção dividida e seletiva), os resultados obtidos sugerem uma atenção dividida fácil sobre partes agrupadas e atenção seletiva difícil em partes dentro de grupos, à medida que aumenta-se a distância. Uma discussão desse trabalho bem como outros trabalhos relativos à percepção de partes em um todo pode ser vista Pomerantz (1986). Pomerantz conclui que o agrupamento de partes em um todo não torna suas partes imperceptíveis ou inacessíveis e que a evidência para um “cola perceptiva” que integra partes juntas é pouca. Segundo Pomerantz quando partes são agrupadas em um todo, novas características emergentes se originam e que estão disponíveis nas partes, mas que são mais salientes perceptivamente e serão atendidas ao invés das partes quando for vantajoso.

3.2–Organização da Percepção em Visão Computacional

A psicologia Gestalt tem ressaltado a importância da organização em visão. Ela demonstrou que as formas possuem algumas propriedades elusivas²¹, imensuráveis, que não aparecem quando analisadas por suas partes constituintes. A Gestalt reconheceu *o que é* computado na organização da percepção mas não explicou de maneira convincente *como* ou *porque*.

Kanisa (1979) apresenta comentários interessantes relativos a conceitos mal interpretados relacionados com a psicologia Gestalt, os quais são importantes quando tenta-se incorporar comportamento Gestalt em computadores. Ele traduz Gestalt como “estrutura

²¹ Indefiníveis e de difícil compreensão, ilusórias.

organizada” em oposição à “forma”, “agregado” (massa) e amontoado (“*heap*”). O conceito global (*whole*) é ordenado, governado por regras e não randômico. Suas outras afirmações são:

- 1) A psicologia Gestalt não é basicamente a psicologia da percepção. A psicologia Gestalt tem sido experimentada em percepção (Koffka, 1935), produção de pensamento (Wertheimer, 1920), resolução de problemas (Duncker, 1945), memória e aprendizagem (Katona, 1940) e muito mais. A teoria Gestalt é um sistema muito genérico de conceitos que podem ser usados para entender virtualmente qualquer aspecto de experiência e comportamento. Assim, esses princípios devem fornecer em muitos aspectos, informações úteis para robôs que exibem comportamento inteligente.
- 2) A psicologia Gestalt propõe que as “propriedades do todo não são o resultado da soma de suas partes” e não o dogma que “o todo é maior que a soma de suas partes”. Assim, uma parte possui propriedades que dependem do todo no qual está incluída. Entretanto, isso não implica necessariamente que tudo depende de tudo mais. A dependência das partes no todo possui uma natureza variável. A psicologia Gestalt simplesmente afirma que as partes do todo não são completamente independentes e assim não podem ser analisadas separadamente.
- 3) A Psicologia Gestalt não rejeita a importância da Atenção, Intenção, Interesse, Atitude, e fatores organizacionais em nossa experiência perceptiva.
- 4) A psicologia Gestalt não rejeita a influência da experiência passada na percepção, entretanto ela rejeita a idéia empírica de que a experiência passada é um princípio explanatório universal. Wertheimer (1923) através da lei da experiência ou familiaridade estabeleceu explicitamente que a experiência passada ajuda na construção das dependências (relações) entre várias partes. Se AB e C, mas não BC têm sido “associados”, então existe uma tendência para ABC ser percebido como AB e C.
- 5) A regularidade de uma figura não é justamente sua simetria. Um dos princípios básicos da psicologia Gestalt é que a organização tende à *prägnanz*, isto é, a tendência de um processo em perceber o estado mais regular, ordenado, estável e balanceado possível em uma dada situação. Não apenas a regularidade geométrica está relacionada mas também a estabilidade, simplicidade e coesão também exercem um papel importante.

É interessante perceber que poucos aspectos da teoria Gestalt citados na figura 2.23 seguir têm sido utilizados em Visão Computacional. Mecanismos de atenção, uso de uma base de conhecimento e outros critérios como regularidade geométrica são raramente consideradas. Existe um aumento crescente no interesse de se incorporar

aspectos perceptivos. Mecanismos de atenção, intenção e atitude são importantes no contexto de tarefas baseadas em Visão Ativa.

Kanisa (1979) sugeriu que a percepção da forma pode ser dividida em dois processos. O processo primário segrega o campo visual em regiões que possuem regularidades espacial e temporal. O processo nos permite ir além da informação dada, é a inferência perceptiva de totalização, *completion*²², de integração, de “preenchimento de lacunas (*gaps*)”, isto é, de inferir aquilo que não está presente. Essa maneira de abordagem é especificamente benéfica no contexto de visão de máquina. O segundo processo poderia auxiliar no preenchimento de lacunas que ocorrem por exemplo por segmentação imperfeita ou em oclusões na cena. Entretanto, como Kanizsa ilustra, explorando transparência poderíamos perceber organizações que são fisicamente impossíveis ou “sem sentido”, baseado em experiência passada mas as quais são percebidas devido a fortes pistas de segmentação fornecida pelo primeiro processo. Isso indica que o segundo processo não determina apenas fatores de organização. Kanisa estudou também o fenômeno de contornos anômalos. Na figura 2.23 podemos um triângulo branco que não foi desenhado. Ele observou que contornos anômalos ocorrem quando suas partes constituintes não estão completas ou auto-suficientes e é necessário para completar partes incompletas da imagem. O fechamento de uma parte é dependente da quantidade de *Prägnanz* que ela possui. Como veremos, muitos esforços têm sido feitos para se quantificar *Prägnanz*.

Figura 2.23 – Contorno subjetivo. A região delimitada pelo contorno subjetivo parece ser mais intensa do que o fundo. Fechamento parece ser suficiente para sustentar a ilusão (Kanizsa, 1979).

²²*Completion*: completamento, completação, aquilo que completa, complemento, fechamento, conclusão, inteireza, acabamento (Michaelis, 1987)

Sarkar & Boyer, (1993) mostraram que a importância da organização pode ser mais explícita, utilizando uma linguagem matemática. Definindo *Casualidade* o evento que um conjunto de características faz parte do (ou é causado pelo) mesmo objeto e *Organização* denota um evento onde as características possuem uma certa organização entre elas. A probabilidade de que as características originam-se da mesma causa comum, dado que tenha sido determinada alguma relação entre elas: $P(\text{Casualidade/Organização})$. Usando-se a regra de Bayes, tem-se:

$$P(\text{Casualidade/Organização}) = \frac{P(\text{Organização/Casualidade}) \cdot P(\text{casualidade})}{P(\text{Organização})} \quad (1)$$

onde $P(\text{casualidade})$ é a probabilidade de que conjuntos são provenientes de uma mesma causa e de alguma maneira captam o *significado* de uma cena. Se considerarmos todo o mundo randômico, essa probabilidade seria baixa mas na vida real é alta. $P(\text{organização})$ é a probabilidade da organização estar presente dentro de um conjunto de características. O valor dessa probabilidade decresce a medida que a complexidade da organização aumenta. Ela capta a probabilidade de ocorrência accidental na organização. $P(\text{organização/casualidade})$ é a probabilidade de organização entre um conjunto de características oriundas da mesma causa. Ela capta o fato de que a matéria é coerente e se comporta de acordo com alguma lei fixa (descobertas ou ainda a serem descobertas). Esse valor é muito alto. Isso leva a inferência de que aquelas organizações que possuem uma probabilidade de ocorrência accidental muito baixa e uma grande probabilidade de ser exibida pela matéria. Dependendo do mundo que possuímos, nosso conjunto de organizações salientes será diferente. Os psicólogos Gestalt estabeleceram um conjunto de propriedades que são importantes na organização perceptiva de imagens em nosso mundo como proximidade, continuidade, similaridade, fechamento e novas propriedades de região comum e conectividade. Essas propriedades são exibidas por qualquer organização saliente e devemos procurar por grupos que exibam tais propriedades. Uma correlação entre o *gap* de separação entre características de um mesmo objeto e o *gap* entre objetos permite coeficientes de correlação significativamente altos. Isso significa que $P(\text{Proximidade/Casualidade})$ é muito alta.

Os tipos de organização consideradas pela comunidade de Visão Computacional são paralelismo, retangularidade e ortogonalidade (formas de simetria), convexidade e continuidade dos contornos das bordas. Padrões exibindo qualquer princípio Gestalt são também significativos. Em resumo, organização proporciona um indicador muito bom da

casualidade comum e a saliência de uma organização depende de sua probabilidade de ocorrência acidental e sua probabilidade de ser parte do objeto.

Os agrupamentos perceptivos diferem uns dos outros com respeito ao tipo de características que eles compreendem as dimensões sob as quais são vistas (Witkin & Tenenbaum, 1983).

É a habilidade da organização imposta aos dados sensoriais que faz a percepção humana tão poderosa e versátil. Esse fenômeno tem sido reconhecido pela Visão Computacional na redução da complexidade computacional do reconhecimento visual. Muitos sistemas de visão organizam primariamente seus dados em bordas ou regiões, em grupos e estruturas perceptivamente significativos e que são invariantes em relação a um grande número de vistas a fim de aumentar a eficiência do modelo de *matching*.

A organização perceptiva pode ser definida como a habilidade de impor organização estrutural aos dados sensorados, de forma a agrupá-los em primitivas originadas por uma causa comum (Sarkar & Boyer, 1993). Esse tipo de organização nos conduz a formar hipóteses com mínimo domínio do conhecimento e entretanto com mínimas restrições, além do fato do nosso mundo não ser visualmente caótico e sim estruturado e organizado.

O papel da organização da percepção está bem ilustrado em um importante experimento relatado por Smith (1987) apud (Sarkar & Boyer, 1993). Um estímulo randômico foi apresentado para um grupo de pessoas que foram requisitadas a reproduzi-lo. O conjunto reproduzido foi apresentado a um outro grupo que foi requisitado a reproduzir o que havia visto. Após aproximadamente doze passos a reprodução final tinha uma estrutura definida. As pessoas envolvidas no teste (gradualmente) impuseram organização a um estímulo randômico. Essa experiência indica a importância da organização na percepção humana. O sistema visual humano valoriza a organização de tal maneira que se esforça para encontrar (ou impor) uma , se não existir alguma organização realmente. Claramente é necessário um comportamento heurístico nesse mundo tão estruturado.

A evolução da *organização (estrutura) da percepção* na Visão Biológica e sua necessidade em sistemas de Visão Computacional basearam-se na característica de que a percepção (extração de significado a partir dos dados sensorados) é um processo inteligente (Sarkar & Boyer, 1993). Isso ocorre particularmente para organismos de ordens superiores e analogamente para modelos computacionais mais complexos. O papel da organização

perceptiva na visão ou em qualquer outra modalidade sensorial promove robustez e eficiência computacional ao processo perceptivo. A importância de se organizar os dados sensorados foi reconhecida especialmente pelos psicólogos da escola Gestalt e seu papel em Visão Computacional foi enfatizado e demonstrado por pesquisadores como Witkin & Tenenbaum (1983) e Lowe (1985). Uma excelente revisão da organização da percepção em Visão Computacional pode ser vista em (Sarkar & Boyer, 1993).

A sofisticação de um sistema de visão depende em grande parte da sofisticação do seu processo de organização da percepção. Nos sistemas biológicos essa capacidade visual é desenvolvida em um grau apropriado para as habilidades modificáveis do organismo, isto é, o grau no qual o sistema visual utiliza para adquirir comida, desviar do perigo ou reestruturar seu ambiente.

A organização da percepção nos permite determinar recursos de maneira inteligente, o que é particularmente importante em sistemas biológicos devido ao custo relativo do tecido neuronal do ponto de vista evolucionário. Como pode ser visto, a organização perceptiva em Visão Computacional utiliza os recursos computacionais eficientemente para extrair organizações a partir das quais levante-se hipóteses sobre características, ao invés de alternativamente aplicar detetores de características sobre toda a imagem.

Uma importante observação feita na revisão apresentada em Sarkar & Boyer, (1993) é que a aplicação da organização perceptiva no sentido Gestalt tem sido mínima ou até mesmo ausente em Visão computacional. Apenas os princípios de proximidade, continuidade e similaridade exercem um papel maior. Embora a importância da simetria tenha sido reconhecida ela raramente tem sido utilizada em conjunção com outros princípios Gestalt. A integração de pistas não geométricas tais como simplicidade e coesão precisam ser exploradas.

4 - CONCLUSÕES

No ambiente natural a redundância de informação permite que a percepção seja suficientemente rápida e verídica para permitir que o observador reaja de maneira rápida e adequada, a não ser no caso de acontecimentos rápidos e inesperados ou que o ambiente esteja limitado por causas naturais por exemplo com iluminação muito reduzida.

É indiscutivelmente verdade que houve muitos progressos no estudo e na compreensão das origens e do desenvolvimento da percepção em crianças. Apesar disso ainda existem muitas falhas em nosso conhecimento. Por exemplo, se a percepção começa com as respostas de unidades celulares nos mecanismos visuais e linhas em diferentes orientações, como é que estas são integradas em padrões e configurações que segundo se reconhece desde os trabalhos dos psicólogos Gestalt, têm significação fundamental na percepção?

Notamos que aparentemente os bebês percebem e diferenciam padrões, no entanto não existem dados quanto às suas respostas aos tipos de configuração estudados pelos psicólogos Gestalt. Será que os bebês realmente percebem mais facilmente as “boas” Gestalts do que as menos “boas”? Aqui, é difícil diferenciar em bebês a tendência para demorar na exploração de padrões que possuem elevado grau de casualidade e incerteza a fim de descobrir sua natureza e estrutura e a possível atração da boa forma como tal.

Quanto a percepção adulta surge uma outra pergunta: será que a bondade no sentido Gestalt pode ser considerado como redundância ou falta de incerteza? Essa equivalência pode ser possível para algumas qualidades por exemplo simetria, regularidade e simplicidade, mas parece improvável quanto a outros aspectos por exemplo continuidade, completude e articulação das partes no todo.

A partir do que sabemos hoje, parece inevitável supor a existências de tendências naturais para perceber as configurações como todos contínuos e coerentes. Estes parecem exercer uma influência tão forte na percepção que provocam a percepção ilusória que é impossível superar. Entretanto, apesar disso, sob certos aspectos a percepção de forma e padrão aumenta sua exatidão com a idade, em consequência da maturação e aprendizagem.

Quando passamos a considerar a percepção de objetos, que constitui a grande massa da nossa percepção diária, torna-se ainda mais difícil estabelecer qualquer relação com a percepção de linhas e até de formas complexas e não-representativas. Não se pode supor que a percepção de objetos seja inata, embora se possa apresentar a hipótese de que o bebê possui uma tendência inata para perceber, não apenas a forma, mas também o ambiente natural através dos objetos e que isso passa a atuar na medida em que se desenvolvem suas experiências com os objetos. Pode-se também supor que o bebê tenha uma tendência natural para perceber um espaço contínuo no qual os objetos estão espacialmente ligados à ele e

entre si. Apesar disso, já sugerimos que a partir da infância a criança constrói integrações complexas ou esquemas, através dos quais o que é percebido se liga à lembranças e conhecimento, sobretudo o obtido através da experiência ativa de manipulação de objetos e movimento no ambiente. Além disso os processos de raciocínio e classificação perceptiva têm influência em tais esquemas de tal forma que a percepção imediata é modificada e corrigida para provocar o aparecimento de impressões mais verídicas do ambiente.

Aparentemente, a percepção e outros processos cognitivos afins se integram em esquemas que se desenvolvem mais em direções específicas de acordo com a experiência, portanto a percepção não apenas tende a ser mais rápida e mais exata em situações significativas mas também estabelece expectativas de tal ordem que a atenção é efetivamente direcionada dirigida para tais situações logo que ocorram.

Apesar disso, a partir da infância atua uma tendência para responder à acontecimentos novos, incoerentes e inesperados, através de mudança rápida da atenção para tais acontecimentos e às vezes passar à ação imediata e adequada. Porém a percepção mais intencional e a inferência podem ser necessárias para a obtenção de compreensão completa da situação, antes da ocorrência da ação. Pouca dúvida pode haver quanto ao fato de que a capacidade para responder a um ambiente em constante mudança é uma das nossas mais fundamentais tendências inatas.²³

O movimento está entre as mais freqüentes variações ambientais. A resposta da atenção ao movimento e a sua percepção aparecem quase que desde o nascimento e parecem inatas. Assim como as formas estáticas, também alguns padrões de movimentos associam à classes de objetos e acontecimentos. Existe uma tendência inata para perceber padrões de movimentos característicos de objetos vivos, onde se incluem os movimentos expressivos das pessoas. A expressão de um sorriso é um exemplo disso. Evidentemente, tais percepções são refinadas e aperfeiçoadas através da experiência.

Já notamos a existência de uma importante interação entre a percepção de forma e movimento, de tal maneira que uma forma que se movimenta em um fundo estacionário se

²³ A inteligência pode ser medida como a capacidade de perceber (compreender) e adaptar-se facilmente às mudanças e está relacionada também com memória.

"Progress, far from consisting in change, depends on retentiveness...Those who cannot remember the past are condemned to fulfill it" George Santayana apud (Sarkar & Boyer., 1993).

salienta e é percebida através da excitação de respostas do mecanismo visual. Em segundo lugar, as partes e aspectos constituintes de uma forma são integrados por seus movimentos entre si. É possível que aqui atue um esquema que coordena as modificações relacionadas às variações das formas quanto à orientação.

Pode haver uma tendência inata para perceber as características espaciais do ambiente. O bebê tem uma disposição inata não apenas para diferenciar entre figura e fundo mas também para atribuir extensão ao fundo. É possível que haja também uma avaliação inata, embora rudimentar da posição do corpo no espaço e principalmente com relação à coordenada de gravitação, entretanto a percepção de tais relações espaciais possa ser adquirida apenas através da aprendizagem.

Os efeitos causados na percepção pela motivação individual e qualidades de personalidade são indiretos e atuam através da aquisição diversificada, por diferentes indivíduos, de determinados conhecimentos e habilidades. Tais conhecimentos e habilidades modificam e até iniciam a organização de esquemas que dirigem a atenção seletiva e a discriminação aperfeiçoada, a identificação e a reação, de acordo com certas orientações, embora sem prejudicar a percepção verídica do ambiente, essencial para a manutenção da vida em um ambiente que muda constantemente e que constantemente nos desafia.

O presente capítulo teve o objetivo de apresentar algumas teorias relacionadas à percepção de forma, entretanto existem muitas outras, algumas estão apresentadas em Loncaric (1998). Procurou-se também estabelecer o papel da psicologia Gestalt no processo de percepção (partes x “todo”) e também sua relação com processos mais complexos como a atenção visual.

CAPÍTULO 3

ATENÇÃO VISUAL HUMANA

1 - INTRODUÇÃO

O problema da atenção visual seletiva é um dos mais antigos na psicologia. Em 1890, William James (1890) fez a seguinte definição de atenção: *“É a posse pela mente, de uma maneira vívida e clara, de um entre vários objetos possíveis simultaneamente ou trens de pensamento”*. *“Vemos que nossa mente é, em cada passo, o palco de possibilidades simultâneas. A consciência consiste na comparação dessas possibilidades entre si, na escolha de algumas e no abandono das demais, graças ao mecanismo inibidor e reforçador da atenção. Os resultados melhores e mais sofisticados são filtrados entre os dados escolhidos pela faculdade imediatamente inferior à esta última, que por sua vez foi selecionada de um conjunto maior de dados mais simples e assim por diante”*.

Outras definições têm sido feitas:

- *“Atenção pode ser definida como um processo cognitivo através do qual seleciona-se informações ambientais (sensações) ou informações processadas ativamente de fontes internas (pensamentos). De uma maneira mais geral, a atenção pode ser definida como a habilidade em focalizar e manter interesse em uma dada idéia ou tarefa, incluindo um gerenciamento das distrações¹”*.

- *“Atenção: ato ou estado de atender especialmente através da mente à um objeto ou evento, ou consciência de pensamento, concentração e esforço mental aplicados à eventos sensoriais ou mentais”²*.

Johnston & Dark, (1986) definem Atenção como o processamento diferencial de fontes simultâneas de informações. Essas fontes podem ser tanto internas (memória e conhecimento) como externas (objetos e eventos do ambiente). Assim, Atenção Visual pode ser vista como uma “cola” que liga essas informações e integra características inicialmente

¹ <http://www.bcp.psych.ualbert.ca/~mike/Pe...reet/Dictionary/contents/A/attention.html>.

² <http://unr.edu:80/homepage/otto/CH5/page2.html>

separáveis em objetos unitários. Sem ela características não podem ser relacionadas umas com as outras (Treisman, 1980).

Segundo Milanese, (1993) a atenção visual é a capacidade de detectar rapidamente regiões de interesse na imagem retinal, a fim de reduzir a quantidade de dados para tarefas de processamento complexas, tais como reconhecimento de objetos.

O domínio da psicologia comportamental adiou as pesquisas dos mecanismos internos da atenção seletiva para a primeira metade deste século. A descoberta de que a integridade da formação reticular do cérebro era necessária para manter o estado de alerta, forneceu alguma realidade anatômica dos mecanismos de excitação (Moruzzi & Magoun, 1949). Essa abordagem levou a muitas descobertas relacionadas aos sistemas transmissores da excitação até as computações relacionadas à seleção de informação.

Os mecanismos de processamento de informação que sustentaram os aspectos seletivos da atenção começaram com os estudos da audição, após a II guerra mundial. Broadbent, (1958) propôs o modelo de um filtro limitado para a informação e localizado entre sistemas sensoriais altamente paralelizados e um sistema perceptivo de capacidade limitada. Experimentos com audição seletiva sugeriram a seleção prévia de mensagens relevantes e perda da informação não selecionada pelo processamento consciente. Estudos fisiológicos sugeriram que o processo de seleção envolveria mecanismos de chaveamento (*gating*) do Tálamo e seria controlado por localizações no córtex pré-frontal (Skinner & Yingling, 1977). Em estudos de processamento de informação humana ficou claro que a informação não atendida era processada pelos níveis corticais mais altos uma vez que interferiam nas informações selecionadas (Posner, 1978). Estudos mais recentes com macacos sugeriram que o Tálamo poderia trabalhar em conjunto com as áreas extra-estriadas para chavear a informação para os níveis mais altos (Moran & Desimone, 1985).

Na década de 70 os psicólogos começaram a distinguir entre processos automáticos e processos controlados (Posner, 1978). A atenção começou a ser vista cada vez menos como um gargalo sensorial e mais como um sistema que fornece prioridade para ações motoras, consciência e memória (Allport, 1980). Esses mecanismos de alto nível da atenção envolvem as áreas frontais e fornecem importantes meios de coordenação entre as atividades cognitivas.

A atenção seletiva é considerada uma combinação de pontos de vista da análise comportamental e computacional, de sistemas cognitivos, arquitetura e circuitos cerebrais e da simulação de redes neurais. Uma suposição básica da abordagem cognitiva neurocientífica é que se pudermos descobrir quais as funções dos neurônios e circuitos durante a atenção seletiva, poderíamos conhecer o algoritmo usado pelo sistema e entender como o cérebro realiza a tarefa de atenção.

As informações do ambiente são úteis ao seres humanos para alcançar determinados objetivos. Para ser eficiente na navegação do mundo físico e em lidar com as propriedades dos objetos, as propriedades da cena visual devem ser percebidas e julgadas. Frequentemente, esse julgamento deve ser rápido. As propriedades essenciais sensoradas que conduzem a um julgamento e consequentemente à ação são: a posição dos objetos, a identidade dos objetos e os atributos dos objetos tais como cor, tamanho, direção e velocidade do movimento. Para se descobrir a identidade e localização do objeto é necessário mais do que simplesmente conduzir a informação ao módulo de julgamento do observador. Antes da informação ser dada ao módulo de processamento especializado (cor, movimento, etc) existem alguns problemas que ocorrem tais como a integração (*binding*) e seleção da informação.

A visão primária (*early vision*) é dita ser *bottom-up*, afetada apenas pela imagem formada na retina e pelas características neurais fixas (*hardware*). A visão superior (*higher vision*) usa os resultados provenientes de computações integrativas, expectativa, atenção, etc para processamento posterior de características salientes e comportamentalmente importantes em uma cena e para a interpretação de seus resultados.

2 – MOTIVOS PARA A EXISTÊNCIA DA ATENÇÃO VISUAL

Quando os seres humanos interagem com o ambiente possuem muito pouco tempo para processar toda a informação disponível. Dessa maneira, apenas uma pequena parte dessa informação (a mais relevante) é selecionada e processada com mais detalhes. Segundo Van de Laar et al, (1997) a relevância da informação depende da tarefa a ser realizada. Esse processo de seleção adaptativo é chamado sistema atencional.

Existem evidências psicofísicas de que pessoas necessitam de atenção visual para realizar várias tarefas visuais. Discussões sobre complexidade computacional sugeriram que seria necessário muito *hardware* para visão de alto nível sem a atenção e que a atenção deve ser um componente fundamental para qualquer sistema de visão realístico (Ahmad, 1991).

No cérebro, os primeiros estágios contém neurônios individuais dedicados a computar características locais em paralelo em cada localização da imagem. Existem neurônios para detectar orientação, cor, movimento, profundidade, frequência espacial, entre outras características (Van Essen & Anderson, 1990). Em termos de tempo computacional isso representa uma estratégia muito eficiente. Devido ao grande número de pixels em imagens realísticas, muitos pesquisadores concordam que um processamento paralelo semelhante é necessário em Visão Computacional, entretanto esse tipo de codificação não é viável em níveis mais altos pois não é possível representar explicitamente cada combinação possível de características em cada posição. Uma estratégia melhor seria ter um processador central para detectar características complexas que possa ser direcionado em qualquer região da imagem. Minsky & Papert, (1969) provou que em geral, Visão não pode ser eficiente realizada com um sistema estritamente *bottom-up*.

Assim, uma outra função importante da atenção é permitir uma interface entre os processamentos de baixo nível e de alto nível. Os primeiros estágios do processamento visual são realizados em paralelo. Devido às razões de complexidade, os estágios posteriores necessitam de um componente sequencial. A interface entre esses dois níveis necessita ser eficiente, flexível e incluir mecanismos para controle *top-down*.

A cada momento quando estamos acordados movemos nossos olhos em uma série de movimentos rápidos conhecidos como *eye saccades* e embora normalmente não estejamos conscientes deles, esses movimentos ocorrem de 5 a 7 vezes por segundo (Ahmad, 1991). Uma vez que a fóvea possui uma resolução muito maior que a periferia, esses movimentos dos olhos nos permite investigar objetos interessantes em qualquer lugar na cena. Se nossa retina pudesse representar cada porção do campo visual com igual resolução, esses movimentos seriam desnecessário, mas para isso necessitaríamos de um número maior de neurônios. Como precisamos economizar tempo de processamento, a evolução ajustou os movimentos dos olhos como um compromisso entre tempo e espaço.

Um outro fato surpreendente é que um fenômeno análogo ocorre dentro do nosso cérebro. Mesmo quando não há movimentos evidentes dos olhos, conhecidos como movimentos sacádicos (*eye saccades*), ou cabeça e do corpo, estamos continuamente prestando atenção à diferentes regiões do campo visual. Experimentos mostraram que essas trocas de atenção “encobertas” (covert) podem ocorrer de 15 a 20 vezes por segundo.

Durante a percepção visual e reconhecimento, os olhos humanos movem-se e fixam-se sucessivamente nas partes mais informativas da imagem, as quais serão processadas com uma resolução maior. Ao mesmo tempo, o mecanismo da atenção visual usa a informação extraída da periferia da retina para selecionar a próxima posição do olho e controlar o movimento dos olhos. Assim, os olhos realizam ativamente uma seleção orientada ao problema e processa a informação do mundo visível sob o controle da atenção visual. Consequentemente a percepção visual pode ser considerada um processo comportamental que não pode ser totalmente entendido dentro do contexto limitado de computações neurais sem levar em consideração os aspectos cognitivos e comportamentais (Rybak et al, 1998).

Do ponto de vista comportamental, forma-se no cérebro uma representação interna (modelo) de novas circunstâncias durante a observação consciente e examinação ativa. A examinação ativa extrai e memoriza relações entre as ações aplicadas e as mudanças resultantes nas informações sensoradas. Um objeto externo torna-se “conhecido” e pode ser reconhecido quando o sistema é capaz de subconscientemente manipular o objeto e prever as reações do objeto à ações aplicadas. De acordo com este paradigma, a representação interna do objeto contém correntes de vestígios alternados nas memórias “sensorial” e “motora”. O cérebro usa essas correntes como “programas comportamentais” no “reconhecimento comportamental” subconsciente. O reconhecimento “comportamental” possui dois estágios básicos: seleção consciente do programa comportamental apropriado e execução subconsciente do programa. A comparação entre os sinais esperados (previstos) com os sinais sensorados após cada ação motora, é um procedimento essenciais na execução do programa.

O argumento prevalecente para explicar porque o cérebro necessita de atenção visual é que existe mecanismo neural insuficiente para lidar com todos os estímulos igualmente. Broadbent, (1971) afirmou “A utilidade óbvia de um sistema de seleção é produzir uma economia no mecanismo. Se fosse realizada uma análise completa até mesmo nas mensagens omitidas, não há razão para seleção”. Devido à essa incapacidade do cérebro

em realizar uma análise completa de todos os estímulos de entrada, a natureza encontrou entre pelos menos três necessidades essenciais competentes: quantidade de informação a ser processada e a que grau, quantidade de volume cerebral a ser devotado à tarefa, rapidez do organismo em responder ao estímulo percebido (Tsotos, 1995).

O principal argumento para explicar a necessidade da Atenção Seletiva é em termos da limitação de processamento do sistema visual, entretanto, como sugerido por (Allport, 1989) a atenção pode ser explicada em um nível mais alto do que análise de característica: existe a necessidade de uma forma de seleção que assegure uma coerência comportamental (atenção para ação). Uma vez que a percepção da visão é um meio que permite as pessoas se interagirem com os objetos que compõem o ambiente (manipular, evitar, etc), um grande número de ações são continuamente extraídas e guiadas pela percepção de objetos. Cada uma dessas ações são controladas por um grande número de parâmetros tais como identidade, posição e aparência dos objetos. O processamento seletivo também é necessário para isolar as informações que definem os parâmetros para ações apropriadas. Se cada processamento de informação gerado por algum estímulo visual, potencialmente conduz a uma ação, é necessário mantê-lo separado dos outros, entretanto eles precisam estar coordenados a fim de completar a atividade e permitir que a atenção possa ser desviada se um ocorrer um novo evento importante.

Segundo Ullman, (1984) o papel da atenção é selecionar localizações de interesse a fim de usá-las como pontos de partida para a aplicação de rotinas visuais e motoras.

Ceccato, (1968) atribui à atenção três funções importantes, entre os vários papéis que desempenha:

- Tornar consciente o funcionamento de outros órgãos táteis, ópticos, acústicos, etc. Por exemplo neste instante acontecem trocas de calor entre nossa roupa e nossa pele, mas isso permanece inconsciente até que concentramos nossa atenção, tornando este fenômeno percebido e seu conteúdo pensado.
- Fragmenta o funcionamento do corpo em unidades, cuja duração varia entre um vigésimo de segundo e um segundo e meio.
- Provoca o aparecimento das categorias mentais (singular, plural, início e fim, parte, resto, inteiro, causas e efeitos, necessário, provável, determinado, livre, etc.) através de várias combinações de estados da atenção (cerca de 2000 combinações).

Segundo Ceccato, (1969) a atenção é uma estrutura que caracteriza o pensamento. “É sempre a atenção que, após ter ativado a consciência, fragmentado e reunido convenientemente em unidades as operações dos vários órgãos, relaciona essas unidades entre si”. Isso acontece através de 150 categorias mentais de relações como ‘e’ ,’ou’, ‘para’, ‘com’, ‘por’, ‘de’, etc”.

Em resumo, sob a ação da atenção, tornamo-nos cientes da existência dos objetos e das idéias, fragmentamos as percepções e combinamos as unidades através de categorias e selecionamos a seguir tudo que for objeto de nossa atenção.

3 – FUNÇÕES DO SISTEMA ATENCIONAL

Não é possível especificar o sistema atencional completo do cérebro mas segundo Posner, (1995) as três funções principais do sistema atencional são: orientação do estímulo sensorial, particularmente localizados no espaço visual, detecção de alvos e manutenção do estado de alerta. Essas três funções realizadas respectivamente pelas redes de orientação, executiva e de alerta serão especificadas à seguir:

- *Rede de orientação*: A rede de orientação possui uma conexão direta com o movimento dos olhos que direciona o estímulo para a fóvea, melhorando a eficiência de processamento dos alvos em termos de acuidade, mas também possibilita a troca de prioridade encobertamente, sem mudar a posição dos olhos ou da cabeça. Essa troca “encoberta” tem a função de guiar os olhos às áreas apropriadas. Existem evidências que as sacadas rápidas necessitam de uma troca de atenção antes que elas ocorram. As áreas envolvidas nesse tipo de atenção são: lóbulo parietal posterior, colículo superior e pulvinar. Danos no cérebro em quaisquer uma dessas áreas reduzem a habilidade da troca de atenção encoberta, entretanto cada área produz uma deficiência diferente, evidenciando que cada área anatômica é responsável por uma operação diferente, mesmo que em conjunto realizem uma função de uma rede. O lóbulo parietal posterior possui a habilidade de desabilitar o presente foco de atenção; o colículo superior é ativado para mover a atenção para a área do alvo e o pulvinar está envolvido em restringir a entrada à essa área indexada. De acordo com Robertson et al., (1988) o lóbulo parietal direito é mais influenciado pelos aspectos globais da figura, enquanto que o lóbulo parietal à esquerda é mais influenciado pelos aspectos locais. Pacientes com lesões parietais possuem dificuldades no reconhecimento de padrões (De

Renzi, 1982). O caminho dorsal que se estende do córtex visual primário até o lóbulo parietal parece mediar a atenção visual seletiva. Dados anatômicos sugerem que o segundo caminho cortical ventral que vai do córtex estriado até o córtex temporal inferior, está envolvido no processamento de cor e forma durante o reconhecimento de padrões (Ungerleider & Mishkin, 1982). Existem evidências também que a rede de orientação está envolvida também em busca visual.

- *Rede executiva*: Existe uma limitação em quanto podemos atender de uma vez. Em percepção, pessoas atendem mais facilmente diferentes aspectos de um mesmo objeto do que o mesmo aspecto em diferentes objetos. Uma das limitações em atender diferentes tarefas ao mesmo tempo consiste na similaridade das informações atendidas. Uma demonstração dessa limitação ocorre quando uma pessoa precisa detectar dois alvos que ocorrem simultaneamente. Nessa situação ocorre uma grande interferência. Essa descoberta fundamenta a idéia de que algum sistema comum está envolvido quando algum sinal (detectado ou acessado na memória) é notado conscientemente. Existe também evidências de que o armazenamento de informações apresentadas recentemente, a geração de idéias provenientes de memória de longa duração e o desenvolvimento de esquemas complexos interferem juntos na detecção de novos sinais (Duncan, 1980). Nesse sentido, o atendimento é determinado pelos eventos do ambiente e objetivos e interesses correntes. Quando apropriadamente balanceados, esses dois tipos de entrada levam à seleção da informação relevante à determinação de objetivos e fornece coerência ao comportamento. O sistema precisa, no entanto, permanecer suficientemente flexível para permitir que os objetivos sejam repriorizados com base nas mudanças ambientais. Esse equilíbrio parece ser afetado por danos no lóbulo frontal.

- *Rede de alerta*. Psicólogos cognitivos vêm estudando mudanças no estado de alerta, em tarefas cansativas com baixa probabilidade de alvo em tarefas de monitoramento em tela de radar militar (Parasuraman & Davies, 1984). Em ambos os casos, há evidência de que o aumento do estado de alerta aumenta a velocidade de detecção do alvo. Pacientes com lesões na área frontal direita possuem dificuldade em manter o estado de alerta (Posner & Petersen, 1990).

A mais convincente demonstração de atenção é o experimento relatado por Desimone & Moran, (1987) que investigou as células do córtex visual de macacos. Eles perceberam que quando os macacos estavam prestando a atenção a uma determinada

localização do espaço visual, as células do córtex temporal inferior (TI) responderam apenas quando os estímulos foram apresentados dentro daquela região. Células dentro do TI estão retinotopicamente organizadas e normalmente respondem a uma grande porção do campo visual. Esse experimento mostrou que o campo receptivo efetivo dessas células pode diminuir para cobrir apenas localizações atendidas.

4 - DEFINIÇÃO DE CONCEITO LIMITE

Mesmo quando a atenção está concentrada em determinado ponto ou determinada área do campo visual, a quantidade que pode ser percebida é limitada. Na realidade, raramente prestamos atenção a acontecimentos isolados. Na maioria dos casos tais acontecimentos são organizados de alguma forma e a percepção como já foi demonstrado, depende da natureza dessa organização. Apesar disso podem surgir situações em que o que pode ser percebido depende da localização das formas ou dos objetos expostos e a atenção pode ser voluntariamente dirigida para diferentes partes do todo.

Segundo Miller, (1962) o fato de nossa mente ser limitada e seletiva, obriga-nos a usar cada símbolo da maneira mais eficaz possível. Quando um problema é demasiadamente complexo para adaptar-se à consciência de um indivíduo, devemos diminuir os vários elementos do problema e escolher os símbolos que os representam. Posto que a consciência é limitada, devemos reunir e simbolizar e a maneira que fazemos isso constitui nosso estilo seletivo, ou seja a maneira de compreender esta fatia do universo.

Examinando mais atentamente como nos referimos a um conceito qualquer (informações armazenadas), podemos perceber que não nos lembramos de todas as características de um objeto. O nosso conceito das coisas é variável e elástico e muda de acordo com a atitude, utilidade, ponto de vista e o compromisso assumido com a atenção. Portanto, o conceito é vez por outra, o efeito de uma escolha entre todas as ligações disponíveis das várias “configurações”. Seja qual for a motivação que nos impele a escolher algumas propriedades de um objeto e seja qual for a atitude (descritiva, científica, ética, estética, antagônica, simpática) com a qual se verifica essa escolha, a mesma mantém vivas, sob o ângulo da motivação e por obra da atenção, somente as características que são tidas como essenciais e inibe as demais. Segundo Miller, (1962) a atenção com a qual examinamos as coisas está evidentemente ligada à possibilidade intrínseca e fisiológica de

recordar simultaneamente todas as propriedades dos objetos e todas as relações que os unem, devido aos limites da própria atenção.

Zingales, (1978) exemplifica o nosso conceito limite de bicicleta, vista sob a motivação “disponibilidade de meios de transporte” e a atitude “descritiva”. A bicicleta nos parece formada apenas de “rodas, pedais e guidão”. Por sua vez a roda parece-nos formada apenas de “pneus, câmeras de ar, arco”. Somente através de um esforço da atenção é que reencontramos para a bicicleta a lembrança de instalação elétrica para o farol e lanterna traseira, quadro, bomba, etc. e para a roda a lembrança de válvula, eixo, raios, etc. Mudando nossa atitude para “científica”, a bicicleta será estruturada em “sistema de propulsão”(por sua vez formado por engrenagem, redução do câmbio, etc), “sistema de locomoção”(rodas, etc), estando o *todo* enquadrado por “princípios de equilíbrio” e semelhantes.

5 - TIPOS DE ATENÇÃO

Existem muitas maneiras nas quais a atenção pode ser classificada (Van der Laar, 1997). A seguir será exemplificado os diversos tipos de atenção:

- atenção seletiva: por exemplo seguir uma única conversa;
- atenção dividida: por exemplo seguir duas conversas simultaneamente;
- atenção exógena: por exemplo um barulho pode desviar a conversa atual em direção à sua fonte;
- atenção endógena: por exemplo, quando a troca da conversação é feita por sua própria intenção;
- atenção encoberta: por exemplo, quando a troca de atenção é feita sem nenhum movimento;
- atenção evidente: por exemplo, quando você se desvia de um orador desinteressante.

A atenção é usada para reduzir a quantidade de informação a ser manipulada por níveis cognitivos mais altos no cérebro, enquanto mantém uma alta performance. A atenção seletiva processa apenas uma fonte de informação, enquanto a atenção dividida processa simultaneamente mais de uma fonte de informação (não necessariamente todas).

Um objeto pode atrair atenção de maneira exógena ou endógena. Objetos que diferem muito de seu ambiente, por exemplo em características tais como cor, tamanho,

continuidade ou orientação tridimensional e objetos que são de relevância especial, tal como nome de alguém, atraem atenção exógena ou *bottom-up*. Na atenção endógena ou *top-down*, os níveis cognitivos mais altos no cérebro influenciam o sistema atencional a selecionar em favor de uma particular característica ou de uma combinação de características. Esta seleção pode basear-se apenas na informação disponível no momento em que a decisão é tomada. Em outras palavras, apenas a informação do objeto que é extraída pré-atentivamente pode ser usada para mudar as preferências do sistema atencional.

A última divisão relevante da atenção é a atenção encoberta e a evidente. A atenção evidente (*overt*) refere-se a movimentos observáveis para trocar o foco da atenção, por exemplo, movimento da cabeça ou dos olhos. A atenção encoberta (*covert*) por outro lado, troca a atenção sem qualquer ação motora.

As investigações neurofisiológicas e anatômicas sobre Atenção Visual têm sido trazidas através de experimentos psicofísicos., isto é, na avaliação da performance humana durante a realização de alguma tarefa visual-motora. Muitas investigações psicofísicas estão preocupadas em estudar a atenção encoberta (*covert attention*) e seus efeitos na realização de tarefas visuais. A avaliação é feita baseando-se no tempo de reação (*reaction time*) ou na detecção de estímulos (porcentagem de erros).

A idéia de que a troca de atenção ocorra internamente , independente do movimento dos olhos exista há cerca de um século (Milanese, 1993). Algumas experiências psicofísicas consistem na discriminação de diferentes itens na imagem, durante um curto período de exposição, onde não é possível a realização de um movimento dos olhos.

No procedimento denominado Busca Visual (*Visual Search*), é apresentado a um observador uma tela contendo múltiplos objetos simples (por exemplo barras), cada um caracterizado por um conjunto de características tais como forma, cor, localização e orientação. O objetivo é determinar o elemento *alvo* de um conjunto de outros elementos denominados elementos “*distrativos*” (*distractors*). O *alvo* pode ser caracterizado por um conjunto pré-definido de propriedades, ou simplesmente ou por possuir uma propriedade diferente dos distrativos. O tempo de resposta é medido em relação à diferentes parâmetros e tarefas (Milanese, 1993).

Uma outra maneira de inferência nos processos de visão é por intermédio do procedimento de pistas (*pre-cuing*). Neste caso, o observador deve discriminar um estímulo periférico sem mover os olhos de um ponto central de fixação. No primeiro caso é dito ao observador que o estímulo irá aparecer em uma determinado lado do campo visual. Em outro caso uma pista tal como uma cruz ou uma luz é amostrada imediatamente antes do estímulo aparecer. A pista tem o efeito de atrair a atenção para aquela posição antes do estímulo real.

Existe muito debate entre os psicólogos e psicofísicos sobre as propriedades do “foco atencional” (Groner & Groner, 1989; Umiltà, 1988). Embora exista ainda controvérsias se o foco de atenção pode aumentar ou encolher, o paradigma mais aceito da atenção é em termos do *zoon-lens model* descrito na próxima seção, no qual o tamanho do foco adapta-se ao estímulo. No caso de vários objetos, um foco circular ou oval encobre todos os objetos. O tempo envolvido na troca da atenção varia de 30 a 250 ms, com uma média de 50 ms, entretanto as trocas de atenção são executadas de 3 a 4 vezes mais rápidas que as sacadas.

Ainda hoje na pesquisa de Atenção Visual, várias questões continuam abertas, como por exemplo a maneira como o processamento da informação visual é direcionado e para onde é direcionado. A primeira questão refere-se ao controle exógeno x endógeno da atenção, enquanto que a segunda questão refere-se a atenção baseada no espaço x atenção baseada no objeto. Como a atenção pode ser direcionada a um objeto ou a uma localização, e de uma maneira *bottom-up* ou *top-down*, recente trabalho apresentado por Lauwereyns (1998) discute a viabilidade da existências desses quatro tipos de atenção. Maringelli & Umiltà (1998) também apresentam algumas experiências relacionadas ao controle da atenção focal. Egly et al. (1994) demonstraram que é mais rápido trocar a atenção dentro de um mesmo objeto do que entre objetos. Chen (1998) também discute este aspecto e argumenta que na troca de atenção dentro do objeto e entre objetos, a organização subjetiva é um fator importante para a manifestação desse efeito

Deubel et al. (1998) demonstrou que a seleção para percepção no caminho ventral e a seleção do alvo através dos movimentos dos olhos estão intimamente ligadas. O processamento visual nos dois caminhos não ocorre de uma maneira puramente automática (*bottom-up*), e sim controlada e dependente da tarefa. Muita pesquisa em psicologia e Neurociência têm sido feita para investigar as propriedades desses processos de seleção (Bundesen, 1990; Desimone & Duncan, 1995; Posner & Petersen, 1990; Schneider, 1993).

Segundo eles, há uma tendência em se medir os efeitos da atenção visual no caminho ventral devido a várias funções sugeridas para a atenção visual como por exemplo:

- Facilitar a detecção (Posner, 1980)
- Permitir integração de características (Treisman & Gelade, 1980)
- Reconhecimento do objeto (Schneider, 1995; La Berge & Brown, 1989)
- Entrada para memória visual de curta duração (Bundesen, 1990)

entretanto, isso não implica que os mecanismos de seleção estejam localizados apenas no caminho ventral (seleção para percepção). Várias teorias sugeriram um papel central do fluxo dorsal em controlar os mecanismos da atenção, muitas vezes denominados “mecanismos de atenção espacial” (seleção para ação). Schneider, (1995) também sugeriu um mecanismo comum para seleção em ambos os caminhos

6 - RELAÇÃO DA ATENÇÃO COM O MOVIMENTO DOS OLHOS

A principal razão biológica para a existência dos movimentos dos olhos é a necessidade de direcionar a alta resolução da fóvea para as partes mais importantes ou “informativas” de uma cena. Essa necessidade é tão forte que origina o fenômeno denominado de campos vazios (*empty fields*), ou seja, quando não há movimento de qualquer tipo, a percepção na retina decai completamente em alguns segundos, através de algum mecanismo de inibição lateral (Yarbus, 1967).

Dodge, (1903) classificou os movimentos oculares, como listados abaixo:

- *Fixação*: Fixação dos olhos em um objeto ou em partes de um objeto, pela fóvea.
- *Movimentos sacádicos*. Consistem de uma rápida rotação da pupila de uma fixação para outra. Esse movimento é uma resposta dos músculos dos olhos a um estímulo na periferia da retina a fim de trazê-lo para o centro da retina (fóvea). Cada movimento sacádico deve ser um movimento de atenção de um objeto ao outro, entretanto a atenção não é apenas seletiva, é também integrativa. A atenção pode ser concentrada mas pode estar também distribuída, assim segundo Gibson, (1986) a atenção de detalhes não é inconsistente com a atenção do *todo*, assim uma série completa de fixações pode ser um único ato de atenção.
- *Movimento de perseguição (tracking)*: É a fixação do olho em um objeto móvel e é muito mais lento do que os movimentos sacádicos.

- *Convergência e divergência*: Convergência é uma rotação interna de cada olho a fim de permitir ambos os olhos se fixarem no mesmo objeto próximo enquanto que divergência permite que ambos os olhos se fixem em um mesmo objeto distante.

- *Movimento compensatório*: É a rotação de cada olho na cabeça mas em posição oposta do movimento da cabeça para compensá-lo.

A atenção encoberta parece estar intimamente relacionada com o movimento dos olhos embora acredita-se que sejam processos separados. Muitas das áreas do cérebro relacionam-se tanto com a atenção encoberta quanto ao movimento dos olhos. Posner et al., (1982) mostraram que a atenção encoberta quase sempre se move para o alvo desejado 200ms antes dos olhos começarem a se mover, embora nem todos os movimentos atencionais resultem em movimento dos olhos. A atenção encoberta é muito mais rápida do que o movimento dos olhos e pode ocorrer 25 vezes por segundo (Julesz & Bergen, 1987) enquanto o movimento dos olhos podem ocorrer de 5 a 7 vezes por segundo. Em termos computacionais o movimentos dos olhos têm um maior custo, faz sentido usar o sistema de atenção encoberta para excluir localizações dos candidatos a alvos para então realizar um movimento dos olhos. Apesar da atenção e do movimento dos olhos terem um relacionamento óbvio, ainda não foi desenvolvido um modelo fisiológico/psicológico detalhado e quantitativo.

Um dos componentes básicos do mecanismo atencional é a tarefa de decidir quais localidades visitar. Existem duas fontes possíveis de informação as quais pode afetar essa decisão: informação *bottom-up* e informação *top-down*.

Informações *bottom-up* referem-se às informações obtidas apenas da imagem. A literatura mostra que vários tipos de pistas podem atrair a atenção. Por exemplo, um objeto contendo uma característica diferente das outras pode atrair a atenção. Yantis & Jonides, (1990) mostraram que estímulos que aparecem repentinamente atraem mais atenção do que estímulos persistentes. Um estudo posterior sugeriu que depois de um tempo, o estímulo que apareceu abruptamente possui a mesma prioridade do estímulo não abrupto sugerindo que a prioridade decai com o tempo.

A informação *top-down* também possui uma grande influência nas localizações atendidas. (Yarbus, 1967) descreve experimentos onde o movimento dos olhos eram altamente dependentes das instruções dadas aos observadores, por exemplo quando

instruídos a descrever expressões nos rostos de pessoas presentes em uma imagem, os observadores se fixavam apenas nas faces e fixação sua atenção nos olhos, bocas e nariz da pessoa. Isso demonstra que é altamente improvável que um sistema visual se fixe em características complexas antes de internamente prestar atenção a elas. Algumas vezes, entretanto, um critério mais simples para guiar as fixações é prestar a atenção em contornos ou regiões de contraste. Podemos concluir que os mecanismos *top-down* (dirigido pelo conhecimento) e *bottom-up* (dirigido pelos dados) são combinados juntos.

No controle *top-down* os pontos de fixação contribuem na definição de um “significado” da imagem e o processo cognitivo segue a análise da informação visual. Um exemplo disso é a análise da sequência de objetos: a duração das fixações é maior para objetos cujos nomes são maiores (Rayner, 1978) apud (Milanese, 1993). Durante a leitura de um texto escrito, os pontos de fixação são também distribuídos de uma maneira ótima, não apenas com respeito ao tamanho da palavra, mas também respeitando-se a estrutura léxica das palavras (O’Regan, 1988) apud (Milanese, 1993).

No controle *top-down* a “composição” da imagem é importante para caracterizar as fixações, por exemplo o tamanho dos objetos e suas posições relativas definem uma organização que é reproduzida pelos caminhos dos olhos. Os olhos geralmente não se movem para regiões uniformes e são atraídos por regiões compactas que definem objetos inteiros e por ângulos e contornos.

Noton & Stark, (1971) desenvolveram um modelo para o movimento dos olhos. O comportamento do observador é distinguido em duas fases, a do aprendizado e a do reconhecimento. O aprendizado ocorre quando o observador está vendo um padrão pela primeira vez e existe a necessidade de memorizá-lo. A fase de reconhecimento ocorre quando a representação armazenada é comparada com o novo padrão de entrada. Noton & Stark observaram que o movimento dos olhos, sob certas condições, que ocorreram durante a fase de aprendizado e a de reconhecimento são os mesmos para o mesmo observador, mas varia entre observadores diferentes.

Um outro problema que ocorre é a integração de características através dos movimentos dos olhos. Quando o nosso olho ou a nossa cabeça se movimenta, a imagem da retina muda. Estudos revelaram que é realizada uma comparação através do conhecimento do movimento do olho (direção e distância) e da medida de similaridade entre as

representações intermediárias da nova imagem da fóvea e as imagens periféricas anteriores (O'Regan, 1988).

Existem algumas evidências de que a atenção encoberta está envolvida na computação de movimento. Cavanagh, (1990) propôs que quando rastreamos um objeto móvel, o seguimos com nossos olhos e cabeça e embora a imagem retinal permaneça aproximadamente constante, o movimento global do objeto pode ser recuperado acumulando-se os comandos motores do olho e da cabeça. Podemos detectar o movimento dos objetos mantendo-se os olhos e a cabeça fixos, dessa forma existem ainda duas possibilidades: a percepção pode ocorrer devido a sinais gerados por detetores de movimento de baixo nível ou se for utilizado no rastreamento um foco móvel de atenção que gera os sinais necessários a percepção

7 - BUSCA VISUAL

A tarefa de busca de um alvo em uma cena visual é natural para as pessoas e nas últimas duas décadas têm sido estudada em detalhes pelos psicólogos. De fato, uma fração significativa do conhecimento psicológico de atenção vem dos experimentos de busca visual. Existe uma forte evidência de que a atenção serial é importante para realizar algumas tarefas de busca, enquanto outras não necessitam desse esforço. Essa dicotomia faz da busca visual um domínio interessante para o estudo de como e porque a atenção é utilizada.

Um objetivo básico da Visão é identificar objetos baseado na combinação de características, por exemplo detectar um círculo-azul (forma-cor). Obviamente pessoas podem associar combinações de características com objetos. Uma vez que objetos podem aparecer em qualquer lugar na imagem, em qualquer tamanho e formato, não é plausível explicitar uma representação em paralelo para cada combinação possível de características em cada localização e escala. Algumas operações podem ser realizadas em paralelo, por exemplo quando um único objeto está presente na imagem pode-se associar a ele uma característica particular que está presente na imagem computando-se um OU (soma) global dos mapas de características, mas quando vários objetos estão presentes poderá haver interferência. Por exemplo se o mapa da característica brilhante global está ativo, o sistema pode não identificar qual objeto é brilhante. Devido a essa interferência, é difícil identificar as localizações individuais dos objetos quando vários objetos estão presentes. Uma solução natural é baseada em um mecanismo que iniba a representação em qualquer lugar exceto na

localização de interesse. Um sistema que usa tal mecanismo pode prestar a atenção a objetos individuais para determinar suas combinações de características.

Entre uma grande variedade de tarefas de busca visual, uma que foi percebida e estudada por Bela Julesz e seus colegas (1983 a,b) é a busca por um alvo que é salientemente diferente dos elementos distrativos vizinhos. Quando a saliência “salta” ao observador mesmo sem uma busca atencional pretendida, esta tarefa é considerada ser realizada pelo sistema visual no estágio de processamento pré-atentivo primário (Bergen & Julesz, 1983 a,b) e (Treisman & Gelade, 1980). Sob essas condições, o tempo de busca é independente do número de distrativos e assim é assumido ser realizada em paralelo para os vários elementos no campo visual. A saliência que produz essa busca característica deriva das primitivas básicas que o sistema visual detecta diretamente (“textons”; Julesz, 1984) ou das características básicas pelas quais o sistema dedica representações específicas (Treisman & Gelade, 1980). As características ou primitivas básicas que “saltam” (*pop-out*) ainda não estão caracterizadas completamente (Wolfe et al. 1992).

Um dos exemplos mais típicos de detecção “sem esforço” é a busca por uma barra de orientação diferente das barras distratoras vizinhas. Essa tarefa possui também um potencial fisiológico conhecido correlacionado com sua performance: uma grande proporção de neurônios na área visual primária são seletivos à orientação de uma barra no campo visual (Hubel & Wiesel, 1968). Nesse estágio de processamento as computações são consideradas principalmente locais (Julesz, 1968). Além disso as respostas neurais nessa área são grandemente determinadas pela estimulação retinal e são muito pouco afetadas pelas influências atencionais extra retiniais relativas às tarefas (Wurtz et al., 1982) e (Haenny & Schiller, 1988). Uma vez que a detecção “*pop-out*” não necessita de fontes intencionais de atenção, o processamento adequado para tarefas pré-atentivas “*pop-out*” têm sido atribuídas ao córtex visual primário (Gallant, 1995). Esse comportamento ao mapeamento anatômico (chamado psico-anatomia por Julesz) foi sustentado pelo trabalho de Merigan et al., (1992) que percebeu que macacos treinados para realizar tarefas “*pop-out*” mantinham essa habilidade mesmo com a área visual secundária (V2) lesionada, ainda que outras tarefas treinadas dependam dessa área.

Uma grande parte da literatura psicológica suporta o argumento acima, por exemplo o trabalho de busca visual de Treisman & Gormican, (1988) que mostra que alvos definidos por uma única característica podem ser detectados em um tempo constante mas os alvos

definidos por uma combinação de duas ou mais características necessita de um tempo linear em função do número total de objetos. Os movimentos dos olhos são excluídos sugerindo que processos internos são responsáveis pelo comportamento seqüencial.

Treisman & Gelade, (1980) sugeriram que a busca por um alvo definido por uma conjunção de características necessita de busca serial enquanto que a busca de características simples não. Em suas experiências, foram mostrados aos observadores figuras semelhantes às mostradas na figura 1 e os mesmos eram requisitados a detectarem a presença de um alvo pré-definido (por exemplo, barra horizontal cinza). Eles perceberam que se o alvo difere de seus distrativos por uma única característica, o tempo de reação não aumenta em função do número total de objetos mas se é necessária a combinação de duas características, então o tempo de reação aumenta linearmente em função do número de objetos. Eles perceberam também que com características conjuntivas o número de tentativas quando o alvo estava ausente foi o dobro de quando o alvo estava presente. Eles concluíram que isso evidencia que o processo de busca termina logo que o alvo é encontrado.

Um outro paradigma de tempo de busca constante e linear foi reportado por Treisman & Gormican, (1988). Eles perceberam que a busca por uma linha oblíqua (18° inclinada) entre várias linhas verticais pode ser feita em um tempo constante, mas a busca por uma linha vertical entre várias linhas oblíquas leva um tempo linear, conforme ilustra a figura 3.1. Essa assimetria pode ser explicada assumindo-se que uma representação anterior (*early vision*) inclui um número finito de orientações que são codificadas por unidades para orientação vertical ou maiores que 18° . Assimétrias semelhantes estão presentes também na detecção de círculos x elipses, linhas simples x pares de linhas, etc. Isso pode ser explicado como busca de presença de uma característica x busca de ausência de uma característica.

Figura 3.1 - Leva-se menos tempo para detectar uma linha oblíqua entre linhas verticais (a) do que detectar uma linha vertical entre linhas oblíquas (Treisman & Gormican, 1980).

Treisman, (1988) testou a busca visual com uma única característica onde os elementos distrativos não são homogêneos. No primeiro caso, os elementos distrativos variaram em uma dimensão relevante (por exemplo busca de um alvo vermelho entre vários objetos coloridos). No segundo caso, os elementos distrativos variaram dimensões irrelevantes (por exemplo, busca de um alvo vermelho entre objetos de diferentes orientações e tamanhos mas com cor constante). O primeiro caso resultou em busca serial e o segundo caso em tempo de busca constante.

Cave & Wolfe, (1990) e Treisman & Sato, (1990) relataram que o tempo de busca pode variar muito entre os observadores. Conjunções de características de cor e forma produzem grandes variâncias.

Treisman & Sato, (1990) e Wolfe et al, (1989) sugeriram modelos onde conjunções podem ser detectadas em um tempo constante usando-se informação *top-down*. Treisman e Sato sugeriram que se as características que não estão presentes no alvo inibem o mapa de prioridades então a localização que contém a conjunção de duas características retém a prioridade mais alta. Wolfe et al, sugeriram um mecanismo análogo usando excitação ao invés de inibição. Pessoas apenas produzem tempo de busca constante para combinações de características muito específicas. Cave & Wolfe, (1990) sugeriram que isso ocorre devido a uma variedade muito grande de ruídos no sistema.

O que se conclui é que certas combinações de características são de fato representadas explicitamente, e que essas combinações de características são aprendidas através da experiência. Shiu & Pashler, (1991) mostrou que o aprendizado de características as vezes depende do estado cognitivo individual e não apenas de entradas retiniais.

Uma outra característica que afeta a busca visual é a organização da cena que possui um impacto dramático no tempo de reação, conforme ilustra a figura 3.2 extraída de Ahmad, (1991). Na figura, os elementos estão agrupados e separados do alvo, dessa maneira, mesmo pensando-se em uma busca conjuntiva, o número de elementos distrativos não afetam o tempo de busca.

Figura 3.2 - Efeito do agrupamento em tarefas de busca visual. Mesmo se tratando de uma busca de conjunção de características. O alvo (barra horizontal cinza) será encontrado em um tempo constante (Ahmad, 1991).

Segundo alguns experimentos a tarefa de associar uma combinação de características com um objeto necessita de atenção. Treisman & Schimdt, (1982) mostraram para algumas pessoas uma imagem contendo dois objetos de diferentes cores e perceberam que quando o tempo de exposição era muito curto, freqüentemente as pessoas trocavam as cores do objeto. Esse efeito, chamado de “conjunção ilusória”, não ocorria quando havia tempo suficiente para o foco. Treisman, (1988) mostrou em outro experimento que a expectativa prévia pode eliminar esses efeitos.

Egeth, Virzi & Gabart, (1984) contestaram a idéia de que para a busca conjuntiva, seria necessário direcionar a atenção serialmente a cada estímulo na imagem. Eles

argumentaram que os observadores apenas restringem a busca quando os objetos possuem a mesma forma ou cor do objeto alvo.

O problema de integração de características sugere que a localização de objetos individuais na presença de múltiplos objetos também pode ser afetada. Wise & Desimone, (1988) mencionou que quando alvos são apresentados simultaneamente o primeiro movimento dos olhos se estabelece em uma posição intermediária.

Algumas experiências mostraram evidências de que a atenção seria necessária para computar relações espaciais (O'Connell & Treisman, 1991). Relações mais complexas necessitam de alguma forma de processamento sequencial, por exemplo determinar em uma imagem contendo círculos e quadrados a localização do quadrado após o segundo maior círculo (Mahoney, 1987).

8 - MEMÓRIA E ATENÇÃO

Outras funções cognitivas podem participar dos processos perceptivos complexos. Neisser, (1967) argumentou que a informação raramente deriva apenas de percepções instantâneas que se apagam imediatamente na consciência; Ao contrário, as impressões se prolongam pelo menos em um curto período na imagem primária da memória. Isso permite a continuidade em nossa percepção do ambiente e também facilita a utilização de lembranças de experiências anteriores e a aplicação de raciocínio e julgamento na avaliação de acontecimentos, antes de reagir ou decidir como agir. Deve-se compreender que, na vida diária, só em momentos relativamente raros, o comportamento é uma função da percepção de estímulos e acontecimentos isolados, mas é determinado pela continuidade global de conhecimento e experiência a que tais percepções estão associadas. A experiência e o conhecimento participam principalmente das inferências que comumente fazemos a respeito da natureza de objetos e acontecimentos. Em muitas situações a percepção imediata pode ser incluída em decisão e julgamento e ser por estes complementada.

Embora a atenção direcionada espacialmente seja um excelente modelo para a investigação dos mecanismos atencionais, os objetos podem ser selecionados baseado em características não espaciais. Recente trabalho indicou que esse tipo de atenção depende, em parte, dos mecanismos neurais para memória, uma vez que o que nós lembramos e

relembramos dos objetos é o que determina sua relevância comportamental, sua familiaridade e novidade, quão bem eles são segregados dos elementos do fundo e assim por diante (Desimone et al., 1995). Quando nós atentivamente buscamos um objeto em particular em uma cena (por exemplo procurar o rosto de um amigo em meio à multidão), é a nossa memória do objeto que guia a nossa atenção. Nesse contexto, “memória de trabalho” e “seleção atenta” pode ser apenas dois aspectos do mesmo mecanismo neural.

Estudos neurofisiológicos têm revelado uma grande faixa de mecanismos neurais que são a base da formação da memória e que podem influenciar a atenção e a competição dos objetos no campo visual, em uma maneira *top-down* e *bottom-up*. Alguns desses mecanismos de memória são explícitos ou cognitivos, enquanto outros são provavelmente implícitos e baseia-se em uma memória não-cognitiva tal como *priming*³ e aprendizagem da percepção. (Desimone, 1992) dividiu esses mecanismos de memória em cinco classes: sintonia (*tuning*), filtragem adaptativa, ativação sustentada, seleção e associação.

A seguir será feita uma breve revisão desses mecanismos em ordem crescente de envolvimento cognitivo:

- *Sintonia*. O mecanismo de sintonia ocorre como resultado de experiências sensoriais ou aprendizagem perceptiva. Neurônios no córtex auditivo, por exemplo, tendem a trocar seus tons preferidos aos estímulos condicionados (Bakin & Weinberg, 1990). Devido ao fato do afinamento modificar os pesos do processamento cortical em direção ao item aprendido, é provável que cause uma modificação nos pesos atencionais. Se praticamos achar um estímulo visual composto por um gradiente de textura em particular por exemplo, o estímulo tende a “saltar” do campo de textura e será em um tempo cada vez menor (Shiffrin & Schneider, 1977).

- *Filtragem adaptativa*. Na filtragem adaptativa, a informação de entrada é filtrada pelos neurônios de acordo com quão similar ela é à informação armazenada tanto na memória de curta duração quanto na de longa duração. Nos macacos, por exemplo, alguns dos neurônios no córtex TI que respondem seletivamente à alguma característica particular do objeto, tal

³ O reconhecimento de objetos e padrões pode ser facilitado se for dado a uma pessoa uma informação antecipada sobre o objeto, chamada de *priming stimulus*. *Priming* pode causar no sistema visual uma representação do objeto na memória. Essa representação pode ser transferida da memória de longa duração onde presumidamente está armazenada, para a memória de curta duração, onde ocorre o reconhecimento. Se o tempo de exposição do *priming stimulus* é muito curto a transferência pode ser incompleta (Leow, 1994).

como uma cor ou forma específica, fornece sua melhor resposta aos objetos que contém aquelas características mas que são novas, inesperadas ou não vistas recentemente (Baylis & Rolls, 1987; Miller et al., 1993). Assim que o novo estímulo torna-se familiar, os pesos sinápticos no córtex se ajustam de tal maneira que a resposta neural é amortecida, um caso de “desprezo à familiaridade” ao nível neuronal. Como resultado, o aumento da familiaridade causa uma “focalização” da atividade no córtex, de tal forma que objetos mais familiares causam uma ativação mais restrita (Miller et al., 1991). Os neurônios que permanecem dentro desse novo foco de ativação são aqueles que fornecem informações mais críticas sobre os objetos. O aumento de atividade dessas células (e das células que recebem suas saídas) a estímulos novos fornecem uma vantagem adicional na competição neural, trocando a atenção para os novos itens presentes na cena. À medida que o organismo aumenta o contato com o item, a memória é reforçada, os pesos sinápticos se ajustam e a ativação do filtro adaptativo decresce, reduzindo sua influência no sistema atencional. Este estará livre para trocar a atenção para o próximo item novo.

A memória pode ser vista como um filtro primário que influencia o processamento do estímulo no campo visual, incluindo sua habilidade de competir para atenção focal. Evidências para segregação pré-atentiva baseada na novidade e familiaridade foi determinado em Wang et al., (1992) que relatou que observadores acharam uma nova letra no meio de letras familiares com um tempo de reação aproximadamente independente do número de itens no campo visual.

- *Seleção*. Enquanto os mecanismos de filtragem podem influenciar o organismo em direção a um novo item, freqüentemente necessitamos suprimir essa orientação a novos itens e atender a um estímulo familiar relevante. Atenção a itens relevantes parece envolver mecanismos de seleção, os quais atuam quando os neurônios estão preparados para responder a um estímulo esperado que é comportalmente relevante.

Em Miller & Desimone, (1994) macacos foram instruídos a sinalizar quando um determinado objeto aparecesse repetido em uma seqüência temporal e a ignorar quando houvesse repetição de estímulos irrelevantes. Uma subpopulação de neurônios no TI que são seletivos a algumas características do objeto tais como forma, textura ou cor, tiveram suas respostas realçadas àqueles objetos que apresentavam tais características e também compararam o item que o animal estava procurando na seqüência. Essas células parecem

estar pré ajustadas ou preparadas a dar uma resposta mais forte à justamente o objeto que o macaco estava procurando.

Esse tipo de células existem no mesmo córtex TI que as células do filtro adaptativo. Nesse tipo de tarefa citada acima, as células do filtro adaptativo estavam suprimidas por qualquer repetição de estímulo na sequência, relevante ou não. Assim, parece existir dois mecanismos paralelos para memória de curta duração no TI: um automático que pode influenciar o organismo em direção à novos ou inesperados estímulos e um outro ativo (“memória de trabalho”) que pode influenciar o organismo a um estímulo particular relevante. A interação entre esses dois mecanismos pode determinar o objeto que será atendido.

- *Ativação sustentada* - Na ativação sustentada, os neurônios são ativados quando um estímulo armazenado previamente é necessário para a memória de trabalho mas os estímulos não são utilizados por longo tempo. Por exemplo, em algumas partes do córtex pré-frontal ventral, os neurônios que respondem à objetos específicos, podem manter-se ativo quando os objetos não estão presentes por muito tempo, contanto que o macaco mantenha o objeto na mente (Chelazzi et al., 1993). Esses mecanismos encontrados no córtex pré-frontal podem prepara os neurônios do TI para objetos visuais esperados. O processo semelhante de manter um objeto na mente está associado com a ativação do córtex pré-frontal que por sua vez influencia neurônios visuais específicos do TI a responderem a um estímulo específico esperado, aumentando as chances dessas células ganharem qualquer competição neural.

Associação - Os mecanismos associativos aparecem como resultado de emparelhamento de diferentes estímulos sensoriais. Se dois estímulos visuais arbitrários ocorrem repetidamente com um curto tempo entre eles, neurônios no córtex TI tende a responder à ambos. Os mecanismos associativos podem interagir com os mecanismos de ativação sustentada. Em Sakai & Miyashita, (1991) quando os macacos foram ensinados que uma imagem de uma maçã é sempre seguida da imagem de uma banana, por exemplo, a apresentação de uma maçã causa a ativação sustentada de alguns neurônios no TI que normalmente respondem apenas à banana. A ocorrência de um estímulo traz à mente o outro estímulo. Isso sugere um papel para os mecanismos associativos e sustentados na lembrança associativa da memória de longa duração.

9 - DESENVOLVIMENTO DA ATENÇÃO VISUAL

Durante muitos anos houve dúvidas quanto a possibilidade de discriminar entre atenção e percepção pois pensava-se que se alguém prestasse a atenção a um estímulo, automaticamente o perceberia. No entanto pesquisas realizadas demonstraram que na realidade pode haver processos de atenção que atuam até certo ponto independentemente dos processos perceptivos.

A formação reticular do tronco cerebral envia impulsos nervosos ao córtex e isso provoca excitação geral, a partir do sono e da passividade, para estimulação sensorial e tais efeitos aparecem em surtos de alta frequência (“potencial despertados”). A formação reticular do Tálamo provoca uma resposta mais persistente e localizada, às vezes denominada “reflexo orientador” em que a atenção é dirigida para tipos específicos de estimulação.

Evidentemente o grau de atenção à estimulação ambiental deve depender em primeiro lugar, da excitação geral pois é esta que determina se o indivíduo está alerta, passivo, sonolento ou dormindo. A percepção de aspectos importantes e significativos do ambiente é mais rápida e exata se o Tálamo atuar de maneira a reforçar os impulsos sensoriais ligados à estimulação por tais aspectos. Fuster, (1958) mostrou que a estimulação direta da formação reticular em macacos aumenta sua rapidez e exatidão na discriminação entre duas formas. Há não apenas a facilitação dos principais respostas perceptivas e acentuação da discriminação, mas também uma supressão dos impulsos não-significativos. Hernandez-Peon et al, (1956) expuseram um gato a uma série de “cliques” e registraram sua resposta auditiva através de eletrodos inseridos no núcleo coclear, no entanto tais respostas cessaram inteiramente quando o gato via um rato ou sentia cheiro de peixe. Processos semelhantes de inibição foram demonstrados no homem (Jouvet, 1957). Os resultados sugerem que no homem a formação reticular atua de modo a despertar o córtex e dirigir a excitação para ele a fim de estimular a atenção a qualquer coisa que tenha significação específica.

Evidentemente antes da ocorrência de qualquer discriminação, a atenção do bebê precisa ser atraída, de forma que observe um objeto e o diferencie de seu fundo. Mesmo na ocasião do nascimento a atenção é despertada e o olhar é atraído por luzes brilhantes e objetos em movimento e um pouco depois, por alguns tipos de desenho. Apesar disso, a

fixação prolongada num objeto não é necessariamente acompanhada por percepção. No entanto o bebê cada vez mais olha à sua volta e examina seu ambiente. Ao fazê-lo, sua atenção é chamada e mantida por determinados estímulos. Com o aumento da idade, tornam-se mais longos os períodos de fixação e os exames visuais de objetos. Existe pouca capacidade para memorização e cada novo aparecimento de estímulo é percebido como acontecimento novo e independente dos outros.

À medida que aumenta a idade e com esta a amplitude da memória, a novidade e a incerteza se tornam cada vez mais eficientes para chamar a atenção. A atenção por estímulos visuais novos e complexos, bem como o interesse por estes também variam com a experiência atual e anterior do bebê.

Dember & Earl, (1957) afirmaram que os acontecimentos que atraem a atenção são:

- Mudança temporal na estimulação, por exemplo estimulação nova, sobretudo na medida em que aquilo que é observado difere do esperado e discorda disto. A novidade e o inesperado do estímulo não devem ser muito grandes pois o observador pode ficar assustado e ser incapaz de responder adequadamente à ele.
- Mudança espacial e heterogeneidade de estimulação, tal como ocorre com material diversificado e complexo. O material não deve ser tão complexo a ponto de estar inteiramente fora de sua compreensão, pois caso contrário irá rejeitá-lo por ser incapaz de responder adequadamente à ele. Todavia essa rejeição pode ocorrer também no caso de crianças pequenas que segundo Piaget podem ignorar aspectos novos de objetos e assimilá-los a esquemas já existentes.

Aparentemente, a excitação e a atenção realizam duas funções um pouco diferentes nesses dois tipos de situação. Na primeira, o indivíduo precisa ser alertado a fim de adaptar-se rapidamente a uma situação nova e inesperada. No segundo caso, precisa preparar-se para adquirir informação quanto à natureza exata do material a fim de identificá-lo.

Berlyne, (1963) verificou que observadores colocados diante de figuras incoerentes da cabeça de um animal ligado ao corpo de outro, olhavam por mais tempo para tais figuras do que para animais comuns. Além disso, quando se apresentavam uma série de figuras coloridas, com formas diferentes e inesperadas incluídas em dois momentos, essas últimas eram observadas por mais tempo do que as séries regulares. No entanto, a repetição dos mesmos padrões provocava habituação e redução das respostas aos estímulos. Em ambientes

normais essa habituação é freqüentemente evitada por mudanças intencionais de visão e por exploração de aspectos novos no ambiente.

Uma outra forma de incoerência na estimulação é que facilmente desperta e atrai a atenção é a intensidade ou grande contraste com o ambiente. Luzes brilhantes e cores fortes tendem a provocar excitação geral e também direção específica da atenção.

É possível que os desenhos mais complexos que apresentam maior quantidade de informação do que os simples exijam exame mais demorado e talvez maior excitação. Apesar disso, existe também uma preferência nítida por certo grau de complexidade pelo fato de ser mais “interessante”. Em Berlyne, (1963) os observadores diziam que os desenhos incoerentes e as figuras complexas e irregulares eram mais interessantes do que as figuras de animais normais e as figuras regulares, embora também menos agradáveis. No entanto é possível que um grau intermediário de complexidade possa ser satisfatório sob os dois aspectos.

Berlyne & Lewis, (1963) encontraram uma interessante interação entre preferência por complexidade e excitação geral e conclui que existe um nível ideal de excitação geral que provoca um grau máximo de atenção e investigação enquanto que a superexcitação pode prejudicar isso.

9.1 – Exploração e Pesquisa

Geralmente não dependemos do que podemos perceber com um único olhar, num período de duração muito pequena. É mais comum que possamos explorar o campo visual, examinando-o com o objetivo de perceber aspectos determinados. Piaget, (1961) admitiu que essa exploração ativa começa na primeira infância, de doze a dezoito meses de idade, quando a criança manipula objetos e varia suas ações anteriores a fim de descobrir o que acontece com eles e aumenta com a idade. Através dessa atividade a percepção imediata é corrigida e refinada pela comparação conjunta de todas as minúcias, a qual chamou de ligação. Piaget percebeu também que os movimentos dos olhos realizados pelas crianças eram mais irregulares, às vezes olhando para fora da figura apresentada e às vezes fixando ou oscilando um ponto por longos períodos. Os adultos fixavam nos pontos significativos

durante tempo suficiente para perceber o que era necessário e depois passavam para o ponto seguinte.

Estudos com busca visual onde um certo número de figuras-alvo deveriam ser identificadas, por exemplo dez triângulos entre diferentes números de quadrados e losangos, mostraram que o tempo necessário para isso aumentava à medida que aumentavam o número de figuras de fundo (Eriksen, 1955) e o mesmo ocorria com o aumento do número de formas que deveriam ser localizadas. Quando se fazia que as formas-alvo fossem detectadas através de qualidades tais como contorno, tamanho, cor e brilho, a identificação de uma única diferença era mais rápida quando a qualidade diferenciada era cor e era mais rápida para forma do que para brilho ou tamanho (Eriksen, 1953).

Boynton & Bush, (1956) mostrou que algumas figuras eram mais facilmente percebidas do que outras. As cruzes eram mais facilmente identificadas corretamente do que triângulos, quadriláteros e pentágonos. Gould & Schaffer, (1965) encontrou diferenças semelhantes na exploração de uma matriz 6x6 que continham dígitos dispostos aos acaso. Embora houvesse uma tendência geral para explorar da esquerda para direita e de cima para baixo, os dígitos 7 eram identificados mais rapidamente, os “2” e “9” mais lentamente, no entanto os erros mais frequentes ocorriam com os dígitos “5” e “9”.

Neisser, (1964) realizou pesquisas sobre processos de busca e verificou que uma letra arredondada, por exemplo “Q” era identificada mais rapidamente quando apresentada apenas com letras de linhas retas. Uma letra de linhas retas, por exemplo “Z” era identificada mais rapidamente quando apresentada apenas com letras arredondadas. O processo de busca se tornava mais rápido com a prática, até que os observadores conseguissem ler dez fileiras por segundo, no entanto levava muito mais tempo para identificar uma linha que não contivesse uma letra diferente. No primeiro caso, adquiriram a capacidade de ignorar todas as letras não-significativas sendo estas percebidas como “fundo”, enquanto que no segundo caso, precisavam ler todas as letras.

9.2 - Predisposição e Expectativa

Em muitos casos a atenção não é despertada apenas como resposta à estimulação decorrente da situação imediata, mas é dirigida de maneira específica por intenções ou

expectativas previamente existentes. Em tais caso, o observador pode estar “predisposto” para perceber determinados aspectos e é possível verificar que está mais preparado para percebê-los do que se isso não tivesse ocorrido.

Em um artigo intitulado “A prontidão perceptiva”, Bruner, (1957) indicou que os padrões estimuladores eram identificados pela classificação dos objetos e acontecimentos em categorias específicas e que essa classificação depende não apenas da semelhança entre as qualidades dos estímulos e as características da categoria, mas também da relativa acessibilidade de diferentes categorias. A acessibilidade era uma função da probabilidade que aprendemos a esperar da ocorrência de exemplos dessa categoria. Por isso estamos mais preparados para perceber objetos ou acontecimentos mais freqüentes e a formular hipóteses que identifiquem os objetos de acordo com essa experiência. No entanto, em estados de motivação e necessidade podemos intencionalmente procurar acontecimentos ou objetos significativos, sem considerar quaisquer probabilidade. As expectativas podem ser despertadas pelo curso de acontecimentos de que participamos no momento ou de informações dadas pelo experimentador. Na medida em que a atenção é dirigida para acontecimentos esperados, ela pode ser afastada de acontecimentos inesperados e pouco importantes.

Usualmente a freqüência de experiências passadas é identificada como familiaridade, supondo-se que quanto mais freqüentemente um acontecimento tenha ocorrido, maior a probabilidade de que se torne familiar. Bruner & Postman, (1949) apresentaram cartas de baralho com as cores invertidas (espadas e paus vermelhos e ouros e copas pretos). Tais cartas eram percebidas com grande dificuldade e às vezes não podiam ser identificadas. Em alguns casos as cores descritas eram as normalmente esperadas, em outros havia uma posição intermediária e o vermelho era descrito como roxo. Portanto as expectativas não garantem necessariamente maior correção na percepção. Para se conseguir aperfeiçoamento, a instrução deve acentuar os aspectos essenciais e os aspectos invariáveis devem ser separados das variações pouco importantes (Gibson, 1966).

As expectativas podem ser criadas de tal forma que os observadores tenham tendência para perceber formas semelhantes às previamente percebidas. A figura 3.3 ilustra uma forma ambígua que foi descrita como “B” por observadores que tinham visto previamente quatro letras maiúsculas e como “13” por pessoas que haviam visto quatro números (Bruner & Minturn, 1955).

Figura 3.3 - Forma ambígua

9.3 - Divisão e flutuação da atenção

Está claro que existem situações em que é difícil verificar a direção da atenção e que o foco da atenção pode variar de um momento para o outro, segundo as circunstâncias. A estimulação que não é significativa para a tarefa principal pode provocar distração que interfere na concentração da tarefa. Essa distração pode ser provocada por súbita interrupção de um estímulo novo, intenso ou pouco usual e tende a atrair a atenção do observador e este está sujeito a abandonar a tarefa em que está empenhado, até que tenha pesquisado a causa da interrupção.

Easterbrook, (1959) verificou que em situações de tensão emocional os observadores tendiam a ignorar alguns aspectos da situação que sob outras condições teriam notado.

Se um observador é solicitado a realizar simultaneamente duas tarefas, a execução da tarefa considerada principal tende a ser menos prejudicada pela divisão de atenção do que a tarefa secundária Webster & Haslerud, (1964), entretanto em alguns casos a tarefa principal pode sofrer considerável influência da execução simultânea de uma tarefa secundária (Gardner & Lohrenz, 1961).

Existem situações em que a atenção parece flutuar entre dois aspectos do material estimulador que não pode ser percebido simultaneamente. Na alternância de figura-fundo, às vezes parte do desenho aparece como figura enquanto a outra parte aparece como fundo e às vezes ocorre inversão, conforma exemplifica a Figura 3.4.

Figura 3.4 - Alternância de figura e fundo

Mesmo em figuras muito simples pode ocorrer alternância. Na figura 3.5 mostrada abaixo, a alternância ocorre entre a cruz horizontal-vertical e a cruz oblíqua (malta). De um modo geral a cruz vertical e horizontal tende a predominar. O fato de orientar um observador a prestar a atenção especificamente a uma determinada cruz prolongava o seu aparecimento mas não suprimia inteiramente a outra.

Figura 3.5 – Cruz de Malta

Alternâncias semelhantes foram demonstradas com figuras significativas como a da “esposa e sogra” mostrada na Figura 3.6 de Boring, (1930).

Figura 3.6 - “Minha esposa e minha sogra” – (Boring, 1930).

Um aspecto pode ser mais persistente do que o outro, seja através da seleção espontânea ou através de estabelecimento anterior ou predisposição. A atenção pode até

certo ponto, ser controlada voluntariamente. Donahue & Griffiths, (1931) verificaram que a alternância era menor com figuras mais complexas e maior quando os dois aspectos possuíam quantidades semelhantes de complexidade e sentido.

É mais difícil atribuir a qualquer efeito de atenção, os fenômenos que ocorrem imediatamente após a outro, embora aqui possa haver interação com os processos de atenção. Já se verificou que a percepção, sobretudo de estimulação de pouca intensidade, demora para desenvolver-se inteiramente. A consciência de estimulação próxima do limiar não é completa até aproximadamente 200 ms depois de seu início (Crawford, 1947). Por isso pode haver interferência entre dois estímulos que se sucedem rapidamente ou até inibição de um pelo outro.

9.4 – Efeitos na variação e mudança na Percepção

Um importante aspecto da relação percebida entre o corpo, seu ambiente espacial e os objetos neste situados é o esquema estável constituído pelo ambiente. Em numerosos fenômenos perceptivos a mudança estimulativa pode ser muito reduzida ou ligada ao esquema espacial. Vernon, (1952) afirmou que “na medida do possível, o indivíduo constrói seu mundo percebido de acordo com a manutenção de máxima estabilidade, permanência e coerência”. No entanto, uma pesquisa experimental posterior demonstrou que, na percepção imediata, a mudança e a variação são essenciais para a manutenção da percepção normal. Estudos realizados por (Hebb, 1949) e Hebb, (1958), mostraram os efeitos de “isolamento” sensorial e perceptivo, como por exemplo a queda da “vigilância” e a tendência para a redução da exatidão da percepção durante longos períodos e das respostas à estimulação monotonamente repetida. Tais situações ocorrem freqüentemente em tarefas repetitivas na indústria como por exemplo inspeção visual, entre outras. A partir de tais estudos ficou claro que um dos efeitos mais evidentes da estimulação invariável é uma redução no estado de alerta e atenção. Portanto, aqui está um tipo de modificação de percepção primária que tem a maior importância nos processos perceptivos mais complexos.

10 – MODELOS TEÓRICOS PARA A ATENÇÃO VISUAL HUMANA

Estudos psicofísicos sobre a detecção e localização de objetos dentro do campo visual sugerem a existência de dois estágios na percepção humana (Neisser, 1967). No

primeiro estágio, chamado pré-atentivo, características simples são extraídas em paralelo realizando-se um processo de segmentação, agrupamento perceptual e levantamento das regiões relevantes que serão processadas posteriormente. A análise de objetos em partes básicas é chamada de processamento *bottom-up* e no sistema visual humano apresenta a característica de não necessitar de atenção consciente. Essas características são extraídas por populações especializadas de receptores que respondem seletivamente. Segundo Treisman, (1987) e Treisman & Gelade, (1980) características tais como cor, tamanho, textura, brilho, orientação, frequência espacial, curvatura e extremidades das linhas são extraídas no estágio pré-atentivo.

O segundo estágio, denominado atento ou focal, utiliza as informações de baixo nível do estágio anterior e também informações de alto nível, tais como conhecimento e contexto, para focalizar o processamento em localizações particulares do campo visual a fim de realizar o reconhecimento e determinar as possíveis relações entre os objetos. Essas relações de alto nível são chamadas de processamento *top-down*, pois partem de um modelo do objeto e descem para o nível das características.

Existem muitas razões para se concordar que a visão aplica analisadores especializados para decompor estímulos em partes e propriedades e que operações extras são necessárias para especificar suas recombinações dentro de um “*todo*”. Em parte a evidência é anatômica e fisiológica. Em particular o esforço para traçar o que acontece com os dados sensorados sugere que os mesmos são processados em diferentes áreas especializadas. Uma área preocupa-se com a orientação de linhas e bordas, outra com cor, outra com direção de movimento. Apenas depois dos processamentos nessas áreas, que os dados alcançam áreas que discriminam objetos complexos (Treisman, 1987).

Existem duas teorias complementares que foram propostas para a Atenção Visual. A teoria baseada no objeto e a teoria baseada no espaço.

A teoria baseada no objeto (Duncan, 1984) sugere que a atenção é sempre localizada em um único objeto ou em um grupo coerente de características. Estímulos que atingem o sistema visual não são vistos como uma mistura de várias cores e brilho, e sim parecem formar grupos de acordo com os princípios Gestalt tais como simetria, compacidade, fechamento e continuidade. Cada grupo de estímulo parece pertencer ao mesmo objeto. Experimentos psicológicos mostraram que seres humanos não apresentaram dificuldades em

identificar duas características de um objeto mostrado em um breve estímulo, entretanto tiveram problemas em identificar simultaneamente as características de dois objetos devido a atenção estar focalizada em apenas um deles.

Por outro lado, a teoria baseada no espaço está restrita apenas a uma pequena região a uma pequena região do espaço conhecida como campo atencional. Resultados experimentais mostraram que um objeto perto do observador são percebidos mais rapidamente do que os distantes (Anderson, 1990). No plano de vista frontal, o objeto próximo ao centro do campo atencional é reconhecido mais facilmente do que os distantes do centro (Eriksen & James, 1986).

10.1 - O modelo da Lente de Zoom para a Atenção (*Zoom –Lens Model*)

Eriksen & James, (1986) se basearam no fato de que as fontes atencionais podem estar distribuídas em toda a imagem ou podem estar concentradas em localizações pré-sugestionadas para criar um modelo de atenção visual chamado de sistema de lentes de *zoom*. Como uma analogia, em seu modelo, lentes de baixo poder corresponderiam a todo o campo, onde as fontes atencionais estariam distribuídas. Um aumento no poder das lentes implica num campo de visão mais restrito, mais também permite que uma alta densidade de fontes atencionais resolvem o problema de discriminação. Em outras palavras, a principal idéia do modelo é que o processamento é sempre realizado em paralelo sobre todo o campo visual ou sobre uma área mais restrita.

10.2 - O modelo do foco de atenção (*Spotlight Model*)

Esse modelo foi proposto primeiramente por Neisser, (1967) e embora ainda cause polêmicas, ele constitui o paradigma mais aceito de Visão Atencional. Ele sugeriu que a percepção pode ocorrer de duas maneiras. A primeira maneira, chamada de processamento pré-atentivo é realizada em paralelo sobre todo o campo visual e é responsável pela segregação de textura e de figuras. A segunda maneira é o processamento atencional, onde pode ocorrer uma análise sequencial de algumas partes da imagem já segregadas e sua ativação só ocorre se o estágio pré-atentivo falhar. Neisser sugeriu que a palavra “análise” não era muito apropriada e deveria ser substituída por “síntese” uma vez que um essa etapa construiria uma representação mais completa dos objetos selecionados. De seu modelo derivaram ainda os modelos de Treisman e de Julesz, que serão abordados abaixo.

10.3 – A teoria dos *textons* de Julesz

Julesz & Bergen, (1983) propuseram um modelo para a percepção de texturas relacionado com os estágios pré-atentivo e focal de Neisser. Seu principal resultado foi o reconhecimento da importância da análise da textura na percepção e na segregação figura/fundo. Sua teoria estabelece que existem elementos perceptuais básicos chamados “*textons*” que são detectados pelos processos pré-atentivos (por exemplo um *texton* pode corresponder a uma barra, uma elipse ou um segmento de reta). Cada *texton* pode ser representados com atributos tais como cor, orientação, etc. A informação usada para discriminar e separar regiões de texturas diferentes não são fornecidas apenas pelos mapas dos *texton* mas também pelos mapas de estatísticas computados sobre ele, como a densidade de *textons*. Julesz sugere que o sistema visual humano realiza essas computações em uma maneira local, evitando esquemas globais. O processamento pré-atentivo não fornece as relações espaciais entre os *textons*, o que seria determinado na visão atenta.

10.4 - Teoria da Integração de características de Treisman (*Feature Integration*)

A teoria de integração de características foi proposta por Treisman & Gelade, (1980) para explicar diferentes resultados no tempo de reação para a busca visual. O modelo usa mapas de características separados que são selecionados por um único foco de atenção.

Em seu modelo, o tempo de reação mostrou dois comportamentos diferentes: constante ou linearmente crescente com o número de objetos distrativos. Alguns alvos parecem ser detectados em paralelo e parecem “saltar” (*pop out*) no campo visual enquanto outros necessitam de uma busca serial entre os itens. O critério que parece determinar o tipo de comportamento é respectivamente se o alvo é definido por uma propriedade (caso disjuntivo) ou por um conjunto de propriedades (caso conjuntivo). O foco de atenção é utilizado quando é necessário detectar uma conjunção de características. O mapa de localizações contém todos os objetos da imagem. O modelo original previa que todas as buscas conjuntivas necessitariam de uma varredura serial e auto-terminável em cada objeto na imagem. Não havia noção de prioridade, assim a busca conjuntiva sempre resultava em tempo de busca linear. O modelo foi desenvolvido a fim de unir vários resultados em um única estrutura. Essa teoria vem sendo modificada a fim de incluir por exemplo noções de prioridades (Treisman & Sato, 1990).

CAPÍTULO 4

ATENÇÃO VISUAL COMPUTACIONAL

1 – INTRODUÇÃO

Um dos problemas principais da Visão Computacional é a capacidade limitada dos sistemas atuais de se estenderem à imagens de tamanho e complexidade crescentes. Embora tenham sido desenvolvidas técnicas eficientes para algumas tarefas de *early vision*, outros problemas tais como agrupamento perceptivo e reconhecimento de objetos ainda representam obstáculos para aplicações em problemas do mundo real. A fim de realizar eficientemente essas tarefas, é necessário estabelecer previamente o tipo da forma do objeto, o número de objetos e/ou a complexidade do fundo (Milanese, 1993). O problema da seleção da informação tem motivado o aparecimento de representações de imagens em multi-resoluções, tal como pirâmides e *wavelets*. As representações em várias escalas são usadas para simular a resolução não-uniforme da retina e permitem estratégias de processamento *coarse-to-fine*. O problema com essa abordagem é que a utilização de filtros passa-baixa ou passa-faixa não extraem informações suficientes da imagem e não podem constituir um modelo de propósito geral para atenção visual.

2 – MODELOS COMPUTACIONAIS

A partir da década de 80, surgiram modelos mais poderosos de atenção que acompanharam o progresso em Neurofisiologia e Psicologia do sistema visual humano. Alguns modelos descritos a seguir representam modelos computacionais da atenção humana e estão diretamente baseados em alguma evidência biológica, enquanto outros sistemas se focalizaram mais nos aspectos computacionais da simulação e objetivam solucionar os problemas práticos com dados reais. A seguir serão mostrados alguns dos principais modelos que vêm sendo desenvolvidos para a atenção visual nos últimos anos.

2.1 – O Mapa de Saliência (*Saliency Map*)

Koch & Ullman, (1985) propuseram um modelo computacional para atenção e que considera muitos fenômenos psicofísicos. No modelo existem mapas computados em paralelo e que representam características elementares tais como orientação, cor, disparidade e direção de movimento. Existem mecanismos de inibição lateral em cada mapa e são extraídas as localizações que diferem significativamente de suas localizações vizinhas. As informações de conspicuidade ou interesse das localizações são unidas em uma única representação chamada de mapa de saliência.

O mecanismo proposto para se selecionar a localização de maior conspicuidade no mapa de saliência foi utilizar uma rede neural do tipo *Winner-Take-All* (WTA). A desvantagem é que a convergência da rede não é garantida e a taxa de convergência depende dos dados. A solução proposta por Koch & Ullman, (1985) foi realizar computações de uma maneira hierárquica, através de uma estratégia piramidal. Na base da pirâmide foram feitas comparações locais em pequenos grupos de células vizinhas e à medida que se sobe nos níveis mais altos as comparações são feitas em distâncias crescentes, até que o valor mais alto seja detectado no alto da pirâmide. As localizações dos vencedores locais são armazenadas em uma outra pirâmide e assim é possível recuperar a localização do vencedor global mandando-se um sinal negativo para as outras unidades. Foram propostos alguns mecanismos para mover o foco de atenção de uma localização para outra, tais como promover o decaimento da atividade das células do mapa de saliência de maneira automática ou através da inibição dos pesos dos mapas de características que produzem o foco atual. Foram sugeridas também que regras de similaridade e proximidade poderiam influenciar a escolha do próximo foco.

(Koch & Ullman, 1985) introduziu a idéia do modelo de pirâmide da atenção e que foi implementado por (Chapman, 1990). Um esboço do modelo pode ser visto na figura 4.1. Os nós em cada nível recebem a ativação dos nós do nível abaixo e transmitem o máximo desses valores para o nível seguinte. O nó no topo da árvore calcula a localização e o valor da localização mais saliente. O sistema implementa um conjunto fixo de rotinas visuais as quais usam a árvore para acessar características da imagem. A principal dificuldade com essa abordagem é que ela não utiliza um foco continuamente variável. Como resultado o sistema não pode lidar com imagens reais mas assume um mecanismo prévio para segmentar os objetos e normalizar seus tamanhos. Assim como o *Guided Search* de (Wolf & Cave, 1994), cada objeto ocupa exatamente uma localização no mapa de saliência. Uma outra dificuldade

do modelo é que o tempo necessário para o foco é logaritmo de acordo com o número máximo de elementos, entretanto esse modelo foi a primeira tentativa concreta de se implementar um sistema atencional.

Figura 4.1 – Pirâmide atencional de Koch e Ullman.

2.2 – O modelo de Chapman

O sistema de atenção visual proposto por Chapman, (1990) foi inspirado pela Teoria da Integração de Características de Treisman & Gelade, (1980) e pela modelo de Koch & Ullman, (1985). A imagem é inicialmente processada a fim de providenciar mapas de características, tais como orientação e cor. Para cada mapa de características são associados *mapas de ativação* a fim de extrair as possíveis localidades de interesse. Cada mapa de ativação atua como um filtro no mapa de características através da detecção de elementos dentro de uma certa faixa de valores. Por exemplo, 4 mapas de ativação podem ser implementados, identificando os elementos da imagem com orientação de 0, 45, 90 e 135 graus. Os mapas de ativação A^k são construídos através do *gating* de todas as localidades do mapa de características F com um AND lógico, seletivo para valores caracterizados pelo índice k (por exemplo 45 graus).

A fim de simular busca paralela, é suficiente computar A através de um OR lógico global de cada mapa de ativação A_k , identificando a presença de elementos dentro de um certo valor. Para a simulação de busca conjuntiva, são necessários mecanismos de seleção. Para essa finalidade Chapman implementou uma estrutura piramidal no topo do mapa de ativação global A como sugerido por (Koch & Uzman, 1985), a qual atua como uma rede WTA efetiva e de baixo custo computacional.

2.3 - O modelo da Busca Guiada (*Guided Search*)

Cave & Wolfe, (1994) desenvolveu a teoria da Busca Guiada (*Guided Search*) para avaliar os resultados obtidos em busca visual conflitantes com a teoria da integração de características. Nesse modelo, um primeiro processo paralelo constrói um conjunto de mapas de características F_k da imagem de entrada, tais como orientação, cor e tamanho. Para cada mapa, calcula-se a probabilidade de cada localidade em conter o alvo. Isso é obtido combinando medidas de interesse *bottom-up* e *top-down*

Uma contribuição significativo do *Guided Search* é esse mecanismo explícito para computar as prioridades usando uma combinação de informações *bottom-up* e *top-down*. A medida *bottom-up* é computada tomando-se a diferença média do valor da característica local com respeito a todas as outras do mapa. Isso proporciona uma medida de conspicuidade (proeminência) para cada localidade. A medida *top-down* depende do conhecimento sobre o alvo a ser achado. Em particular, o conhecimento sobre a eficiência de diferentes mapas de características para revelar o alvo é usado para ponderar a diferença do valor real da característica com o valor esperado. Esses dois componentes são descritos pela equação 4.1:

$$A_{x,y}^k = e^{\left(\sum_{j=0}^n \frac{|F_{xy}^k - F_{uv}^k|}{n-1} \right)} - d_k |F_{xy}^k - t_k| \quad (1)$$

onde d_k é a eficiência da característica k em descobrir o alvo, e t_k é o valor esperado para o alvo. Os índices espaciais (u,v) na soma varia sobre toda as localidades da imagem.

A operação do *Guided Search* é ilustrada na figura 4.2. Para cada dimensão de característica, o estágio paralelo determina os elementos que mais se aproximam do valor do

alvo para aquela dimensão e que difere dos outros elementos do campo visual. A informação de cada dimensão de característica é somada para cada elemento a fim de produzir um mapa de ativação global, que possui um valor que representa a probabilidade do alvo em estar localizado naquela posição. O estágio serial processa o elemento de maior ativação no mapa de ativação. Se o elemento em consideração for considerado não ser o alvo, ele é eliminado para futuras considerações. Considerando todos esses elementos em ordem decrescente de semelhança, ou determina-se o alvo rapidamente, ou pode-se concluir que o alvo está ausente, sem ter que considerar, como no modelo original de Treisman, todos os elementos em uma ordem aleatória.

Figura 4.2 – Modelo de Busca Guiada de Cave & Wolfe

Nesse diagrama cada mapa de característica foi dividido em dois mapas para mostrar os efeitos *bottom-up* e *top-down* separadamente. Cada localização em um mapa de característica apresenta uma tonalidade de acordo com a sua ativação, onde o tom mais claro representa uma ativação mais alta.

O modelo assume que os objetos já estão segmentados e ocupam uma localização fixa por objeto. Uma desvantagem do método é que a comparação de cada localização com umas com as outras necessita de $f.n^2$ conexões onde f é o número de mapas de características e n o número máximo de elementos. Finalmente, o modelo é muito sensível à ruídos

É interessante notar que, embora esse modelo seja computacional, ele excede a performance humana em muitos problemas de busca conjuntiva. Realmente, se o valor de ativação pode ser assumido a ser uma medida perfeita de semelhança, o estágio serial teria apenas que pegar o elemento de valor mais alto no mapa de ativação global A , e esse poderia corresponder ao alvo na maioria dos casos. Afim de reproduzir a performance humana mais aproximadamente, deve ser adicionado ruído no mapa de ativação. Dessa maneira, uma certa quantidade de incerteza nas medidas A_k é simulada, e um certo número de esforços deve ser realizado pelo estágio sequencial a fim de se detectar o alvo.

2.4 - O VISIT de Ahmad

De uma maneira semelhante ao modelo de Chapman, (1990) a arquitetura proposta por Ahmad, (1991) para modelamento de busca visual é baseada em uma primeira decomposição de imagens de entrada em um número de mapas de características (vermelho, verde, vertical e horizontal). Como no sistema de Chapman, a atividade em cada mapa de característica é detectada por um OU lógico global. Os mapas de características são processados pela rede de *gating* e pela rede de prioridade, conforme ilustra a figura 4.3. A rede de *gating* recebe as coordenadas da região atendida e inibe todas as localidades remanescentes do mapa de características, implementando assim o foco e armazenando em um mapa (*gated map*) apenas as características selecionadas. Um sistema de reconhecimento de alto-nível essencial poderia apenas acessar essas características, evitando interferências com as informações relacionadas com outros objetos.

A rede de prioridade objetiva ordenar as localizações da imagem de acordo com sua relevância. A relevância de uma localidade é computada com base em diferentes pesos W_k , $k=\{\text{vermelho, azul, horizontal, vertical}\}$ ajustado pelo usuário para os mapas de características. No modelo assume-se que o *objeto alvo* é definido, e que as características que o caracterizam são conhecidas previamente. Os valores da rede de prioridade $A_{x,y}$ caracterizando a relevância da localidade (x,y) são definidos pela equação 4.2:

$$A_{x,y} = \delta \left(\sum_{u,v} \sum_k w_k \cdot F_{u,v}^k \right) \quad (2)$$

onde $\delta()$ é uma função sigmoideal, e a soma é estendida a todos os pixels (u,v) pertencendo ao campo receptivo centrado em (x,y) . Isso corresponde a computar a soma ponderada de todos os mapas de características e aplicar um *threshold* aos resultados.

Figura 4.3 – Visão Geral do VISIT para busca visual.

Esse modelo é interessante pois simula a busca paralela e serial e sua implementação foi especificada apenas em termos de redes neurais artificiais, entretanto, similarmente aos modelos anteriores, ela tem sido aplicada apenas em dados sintéticos, e muitos detalhes são perdidos para ela ser aplicável a imagens reais.

2.5 - O Neocognitron

O Neocognitron de Fukushima é provavelmente a primeira rede neural que integra o conceito de atenção Fukushima, (1986) embora seu modelo não estivesse diretamente motivado em bases biológicas.

Fukushima, (1986) descreve um modelo que combina uma rede hierárquica e relaxação iterativa. O nível de saída da rede contém uma unidade para cada padrão a ser

reconhecido. Dada uma imagem, os nós de saídas dos diferentes padrões respondem com um certo grau de atividade. A unidade com a resposta maior é selecionada e um sinal é transmitido de volta através dos caminhos ativados pelo padrão de entrada, distinguindo a detecção desse padrão. Os caminhos correspondentes aos outros objetos são atenuados gradativamente na ausência desse *feedback* positivo. Uma propriedade desse modelo é que a forma do foco de atenção depende apenas da forma do padrão de entrada, entretanto o processo é bastante lento. Esse modelo tem sido utilizado no reconhecimento de caracteres Kanji chineses (Fukushima et al., 1991).

O problema encarado por Fukushima é o reconhecimento de múltiplos objetos, possivelmente sobrepostos em uma imagem, e para tolerar pequenas distorções na forma e tamanho. O exemplo típico é uma imagem de caracteres. A rede é capaz de prestar atenção seletiva a cada caracter de cada vez, e então reconhecê-los individualmente. A imagem de entrada é alimentada para as unidades da camada inferior da rede hierárquica de múltiplas camadas, cujas unidades de saída corresponde a objetos de diferentes classes (caracteres).

O funcionamento da rede baseia-se em aprendizado iterativo, competitivo e auto-organizado. Para qualquer imagem, todas as unidades de saída iram responder com alguma atividade. A unidade com a resposta mais alta será selecionada, e um sinal de reforço é enviado para as células conectadas nas camadas subjacentes. Dessa maneira, todas as células que contribuíram para o vencedor irão receber um sinal positivo de retorno, enquanto a atividade em todas as células remanescentes irá decair. A consequência é que atividade da parte da imagem que não contém o objeto vencedor é atenuada, as interferências são reduzidas e a recuperação do objeto pode ser refinada adaptando os parâmetros das células de extração de características do sinal selecionado. Dessa maneira o sistema presta atenção a cada um dos padrões por vez.

O sistema pode deslocar atenção aos padrões remanescentes de uma maneira automática. Uma vez reconhecido o objeto, o sinal auxiliar, que foi enviado às unidades que contribuíram para isso, é interrompido. Um termo de decaimento causa um decréscimo em suas atividades. Dessa maneira, as células inibidas anteriormente podem recuperar seus ganhos, permitindo que um novo padrão seja reconhecido.

Uma vantagem desse modelo é que a “janela de atenção” adapta-se a forma do padrão de entrada a ser reconhecido, e pode assim ser arbitrário, permitindo que objetos escondidos parcialmente sejam detectados através de grupos de regiões não-conectadas de

atenção. Entretanto, a rede é muito complexa, e a propagação da atividade tem que ir através de muitas camadas a cada iteração. Por essa razão a convergência para cada região de atenção é muito devagar.

2.6 - O modelo de Mozer

Em contraste com os modelos baseados em pirâmide, Mozer, (1988) e Mozer, (1991) descreve um modelo para reconhecimento baseado em relaxação iterativa, e modela muitos dos efeitos perceptivos associados à leitura. Em seu modelo a região de atividade na camada retinotópica permite a passagem da atividade dos mapas de características para os níveis mais altos. O foco corrente é determinado por uma regra iterativa. A cada passo as unidades comparam suas atividades com as unidades vizinhas e ajustam seus valores de acordo com uma regra de atualização competitiva localmente. A rede seleciona uma única região de máxima atividade. A principal vantagem desse modelo em relação aos piramidais é uma região continuamente variável. O modelo também incorpora a habilidade de ponderar mapas de características individuais, embora essa capacidade não ter sido usada nas simulações.

2.7 – O modelo de Sandon

Sandon, (1990) descreve um modelo que utiliza também uma regra iterativa mas realiza a computação em várias escalas espaciais simultaneamente. Um dos problemas com os modelos iterativos é que o tempo de estabilização é muito sensível ao tamanho e tipo da imagem. Esse tempo pode ser muito alto se existirem regiões de atividades similares que estão muito separadas. Um outro aspecto é que a seleção da região é completamente baseada na informação local, não existindo um mecanismo explícito que incorpora informação *top-down* para deslocar a atenção.

2.8 – O modelo de Milanese

O modelo proposto por Milanese (1993) é um modelo puramente *bottom-up* e está estruturado em três estágios principais. Primeiro, são extraídos mapas de características da imagem (como por exemplo orientação, curvatura, contraste, etc.). O segundo estágio

consiste em se determinar um número correspondente de mapas de conspicuidade (através de um banco de filtros de várias escalas e de várias orientações), os quais enriquecem as regiões de interesse em cada mapa e por último é utilizado um processo de relaxação não linear, implementando um esquema de interação competitiva e cooperativa, para integrar os mapas de características em uma única representação. Uma das principais contribuições desse trabalho é poder ser aplicado em imagens reais, complexas e não segmentadas.

3 – CONCLUSÕES

Na literatura da atenção visual seletiva, muitos modelos foram propostos. Por exemplo a teoria da integração de características - Treisman & Gelade, (1980), SLAM - Phaf et al, (1990), VISIT - Ahmad, (1991), *Neocognitron* - Fukushima & Imagawa, (1993), Teoria da Atenção de Bundesen (1990), *Dynamic routing circuits* - Olshausen et al, (1993) e o *Guided Search* - Wolfe, (1994), cada um abordando uma metodologia diferente, como pode ser exemplificado nos modelos iterativos, em pirâmide e conexionista apresentados. Uma excelente revisão dos principais modelos de atenção pode ser vista em Milanese (1993).

A idéia central no estudo da cognição visual primária é que o processamento visual é feito em dois estágios: um processo primário opera em paralelo sobre todo o campo visual e um processo posterior que processa apenas a informação de uma parte limitada do campo visual de cada vez. Essa idéia apareceu na psicologia cognitiva (Neisser, 1967) e formou os fundamentos para a Teoria da Integração de Características de Treisman (Treisman, 1986; Treisman & Gelade, 1980) e a Teoria dos Textons de Julesz (Julesz & Bergen, 1983). Embora essas teorias empreguem terminologias diferentes, ambas empregam o conceito de um estágio paralelo que identifica as características básicas que estão presentes e um estágio serial mais lento combina (integra) essas características a fim de produzir representações complexas dos objetos visuais.

Julesz e Treisman apresentaram um grande quantidade de dados que suportam as teorias de integração de características e seus resultados têm sido expandidos por diversos outros autores. Muitas experiências realizadas utilizaram o paradigma de Busca Visual, no qual consiste na busca de um elemento alvo em meio a elementos distrativos. Quando um alvo difere de todos os outros distrativos por apenas uma característica, as teorias de integração de características sustentam que o estágio paralelo deve ser capaz de detectá-lo e

assim achar o alvo rapidamente, independente do número de distrativos (Tresiman & Gelade, 1980). Essas teorias afirmam no entanto que em caso de conjunção de características, onde o alvo compartilha no mínimo uma característica com os elementos distrativos, o estágio serial deve processar cada elemento individualmente até que o alvo (combinação) seja determinado e necessita de mais tempo à medida que o número de elementos na cena aumentam. Existem entretanto muitas questões abertas em relação como o sistema visual humano faz essa integração de características.

Cave & Wolfe, (1990) desenvolveram uma outra teoria para integração de características onde o estágio paralelo (busca de características) guia o estágio serial (conjunção de características) e escolhe quais os objetos devem ser processados. A simulação em computador do *Guided Search* produziu os mesmos padrões gerais produzidos pelos seres humanos sob investigação, em diversas tarefas visuais. Segundo Cave & Wolfe, o estágio paralelo deve ser capaz de usar uma característica simples para guiar a busca serial, excluindo os elementos com valores errados e direcionando a busca para os valores corretos naquela dimensão. Essa modificação preserva no entanto a idéia essencial de que as dimensões das características são processadas de maneira independente no estágio paralelo.

Muitos conflitos originaram-se quando foram utilizadas características diferentes em buscas conjuntivas. Nakayama & Silverman (1986) mostraram que para alguns tipos de conjunção de características o tempo de reação permaneceu constante à medida que o tamanho do conjunto de objetos aumentou. Os alvos nessas buscas paralelas são geralmente definidos por uma conjunção que envolve tamanho ou profundidade estereoscópica. Outros problemas surgiram do estudo apresentado por Wolfe et al. (1989). Em uma conjunção de características de cor e forma algumas pessoas uma função de resposta característica de busca serial, enquanto que outras pessoas envolvidas na experiência responderam muito rápido independente do número de características. Esses novos dados sugerem que as diferenças entre os estágios de busca paralela e serial pode não ser tão bem definido quando o alvo está envolvido em um conjunção de características. Essa variação em performance entre as pessoas investigadas sugere que existe algum fator que varia de indivíduo para indivíduo e que pode fazer uma diferença entre conjunção serial e conjunção quase paralela.

Em grande parte, as teorias de integração de características fazem sentido computacionalmente. Uma vez que as comparações entre as características são feitas apenas dentro das dimensões e apenas para elementos dentro de uma determinada distância, as operações são simples e podem ser facilmente realizadas localmente e em paralelo sobre o

campo visual. A implementação dessas comparações de características em *hardware* é muito plausível, entretanto a conjunção de características em paralelo apresenta um maior custo computacional pois necessita de um hardware dedicado a cada combinação possível de características e em cada localização. Nesse sentido, as teorias de integração são plausíveis pois permitem que um conjunto de *hardware* de conjunção possa ser usado em muitas localizações diferentes. Em geral, os mecanismos dessas teorias também realizam uma função visual muito útil que é a identificação de características visuais potencialmente importantes de maneira rápida e eficiente.

De maneira análoga às outras teorias, o *Guided Search* possui um estágio paralelo que é sensível as diferenças de características e um estágio serial que realiza operações mais complexas. A grande diferença consiste na interação entre os estágios paralelo e serial. Nas Teorias de Integração de Características, o estágio paralelo tenta encontrar uma alvo característico. No *Guided Search* o estágio serial pode determinar o alvo mais rapidamente porque utiliza a informação proveniente de todas as dimensões para determinar quão parecido é cada elemento com o alvo. A informação fornecida pelo estágio paralelo nem sempre é perfeita e portanto o alvo nem sempre é determinado imediatamente.

Na conjunção de característica padrão, o alvo é o único elemento que será identificado como alvo potencial em cada dimensão de característica. Devido o estágio paralelo do *Guided Search* combinar as informações de todas as dimensões de características é possível determinar uma conjunção alvo imediatamente, entretanto essa determinação não é perfeita pois está sujeita a uma certa quantidade de ruído e em algumas tentativas é possível que algum elemento distrativo pode ser identificado como sendo mais promissor que o alvo.

A maneira na qual o estágio paralelo guia o estágio serial é semelhante ao modelo de dois estágios proposto por Hoffman (1978, 1979). Hoffman motivou-se com alguns dados que mostraram que alguns pessoas, em uma tarefa de busca visual, podiam rapidamente excluir alguns alvos, mas necessitavam de um período relativamente longo de tempo para identificar o alvo. Ele propôs um estágio paralelo para encontrar um estímulo similar ao alvo e um estágio serial que compara os estímulos com cada um dos alvos. O modelo *Guided Search* demonstrou vantagens em se combinar essa idéia com os argumentos das teorias de integração de características quanto a detecção de características em paralelo e a integração de características serial. *Guided Search* também adicionou a essa combinação outras duas idéias. Primeiro, embora o estágio paralelo não possa representar todas as combinações

possíveis de características em cada localização, ele pode combinar informações de diferentes dimensões de características para medir quão semelhante o estímulo é do alvo. Segundo, a informação produzida pelo estágio paralelo não é perfeita e variações nessa qualidade entre as pessoas investigadas podem levar a uma grande diferença individual de performance.

Assim como as teorias de integração de características, *Guided Search* utiliza uma representação separada do campo visual para cada dimensão de característica. Nesse modelo, cada uma das representações é um mapa de característica simples, onde cada localização contém um valor de ativação. Dessa forma, cada localização no campo visual possui uma ativação para cor, outra para orientação e outra para cada dimensão de característica restante.

A componente *bottom-up* dentro de cada mapa de características não depende do valor do alvo esperado para aquela dimensão e será o mesmo quando não se conhece nada a respeito do alvo. Cave & Wolf assumem que essa é a base para o “*pop-out*” visual (Treisman, 1986). Para uma localização em particular, dentro de um mapa em particular, a componente *bottom-up* baseia-se na diferença de valor para aquela característica naquela localização e o valor para cada outra localização na cena. Dessa maneira, se todos os elementos na cena possuem a mesma cor, todas as serão zero e a componente *bottom-up* de ativação de cor será baixa.

O trabalho apresentado nesta tese inspirou-se principalmente no modelo de Cave & Wolfe, que também baseia-se em objetos segmentados e utiliza uma representação *bottom-up* que mede a conspicuidade entre características independentes. Neste trabalho entretanto, através da lógica *fuzzy* e de suas regras de inferência, foi possível fazer uma análise global da conspicuidade dependente do contexto em que as características aparecem. Assim, um objeto que apresenta uma alta conspicuidade de cor e tamanho em relação aos outros objetos, pode apresentar um índice de saliência maior do que um objeto que apresenta uma alta conspicuidade de forma ou orientação. Neste trabalho a informação *top-down* é integrada de maneira diferente do *Guided Search*, entretanto as características do alvo também se baseiam no conhecimento prévio da tarefa a ser realizada.

PARTE 2

O MODELO PROPOSTO

CAPÍTULO 5

ESTRATÉGIA ATENCIONAL PARA BUSCA VISUAL E RECONHECIMENTO INVARIANTE DE OBJETOS ATRAVÉS DA INTEGRAÇÃO DE CARACTERÍSTICAS *BOTTOM-UP* E *TOP-DOWN*

1 - VISÃO GERAL DO MODELO PROPOSTO

O objetivo deste trabalho é apresentar uma metodologia para detectar regiões de interesse em uma imagem, baseando-se em informações fornecidas pela Teoria dos Momentos, tais como orientação, tamanho, formato, posição (distância) e também informações de média de nível de cinza, realizando-se o reconhecimento dos objetos somente quando isto for necessário para a realização de uma tarefa específica.

O trabalho foi inspirado nos mecanismos de foco e seleção da atenção visual humana embora seja importante enfatizar que o objetivo não é construir um modelo biológico, pois muitos aspectos da atenção visual humana ainda não estão totalmente claros e muito estudo deve ser realizado ainda para a construção de um modelo realmente plausível. Imagens contendo objetos geométricos simples e em um fundo uniforme são usadas, mas também será avaliada a performance do sistema em imagens reais, adquiridas por uma câmara, contendo objetos mais complexos e ruidosos.

Neste trabalho, o papel do sistema atencional não é realizar algum tipo de segregação de objeto, mas discriminar entre objetos e selecionar o mais “interessante” entre eles de acordo com uma tarefa específica. Outro papel da atenção visual explorada neste trabalho é o de processar e combinar diferentes fontes de informação independentes e simultâneas. Conforme relatado nos capítulos anteriores, a Atenção Visual pode ser vista como uma cola que integra diferentes fontes de informações internas (memória e conhecimento) e fontes de informação externas (objetos e eventos). Sem a atenção, características não podem ser relacionadas umas com as outras.

A utilização de estratégias atencionais em visão permite muitos benefícios. Talvez a aplicação mais imediata é no contexto da visão ativa, onde é possível a mudança proposital dos parâmetros de aquisição, conduzindo a uma redução na complexidade. Câmeras móveis controladas por computador podem ser capazes de extrair uma sequência de localizações de interesse na imagem. Uma outra classe de problemas que se beneficia com sistemas de atenção é o de supervisão. Existem muitas aplicações cujo objetivo é a detecção de configurações “estranhas” de padrões podem estar relacionados com situações perigosas ou de insegurança. Embora ainda não seja possível um modelo exato, os sistemas atencionais devem ser capazes de detectar situações de interesse e acionar a ação de um supervisor ou rotinas mais especializadas. Uma aplicação de alto impacto da atenção visual é a inspeção e controle de qualidade. Existem muitas aplicações diretas em problemas industriais para controle de qualidade de objetos manufaturados. Existe por exemplo, uma grande demanda em sistemas automotivos em rejeitar os objetos defeituosos ou agrupar em diversas classes. Dentro desse contexto, um objeto defeituoso pode atrair a atenção porque viola alguma regra de “regularidade”. Uma outra classe potencial de problemas industriais é a seleção e ordenação dos objetos de interesse. A capacidade de selecionar regiões de interesse pode simplificar passos como a comparação com modelos armazenados. Em todos esses casos, está claro que um sistema não pode ter um controle puramente *top-down*.

Neste trabalho, inicialmente são extraídas regiões uniformes através de um algoritmo de “crescimento de região”. O sistema atencional extrai então atributos para os objetos dentro do foco e é computada uma medida de interesse para cada localização baseada inicialmente em uma medida de conspicuidade *bottom-up* entre as características das regiões. A componente *top-down* pode ser integrada de diversas maneiras, variando de acordo com a aplicação ou através por exemplo da especificação de *templates* dos objetos de interesse. Neste trabalho, a informação *top-down* refere-se a características esperadas para o objeto alvo e uma medida de interesse é calculada de acordo com a relevância das características na realização de uma tarefa específica. Isso permite reformular o problema de reconhecimento em termos de *indexação* visual, isto é, através da escolha de primitivas selecionadas que são diretamente usadas para ativar hipóteses do objeto na base de dados dos modelos sem a necessidade de se construir uma representação completa do objeto.

Existe uma maneira dos sistemas tradicionais de visão explorar as habilidades dos sistemas atentos, sem ter que processar toda a imagem de uma vez. Uma solução pode ser definir uma pequena janela de atenção que pode ser deslocada por toda a imagem. O sistema de reconhecimento do objeto e o sistema de atenção seleciona uma região de interesse e a

janela pode ser deslocada e centralizada em novas regiões. No contexto dos sistemas de visão tradicionais, as regiões de atenção podem ser usadas para o módulo de reconhecimento do objeto para restringir a informação na imagem de entrada. Assim, uma nova imagem diferente é obtida após uma troca de atenção. Essa nova imagem contém partes (ou informações) de uma cena que não era conhecida antes, entretanto as trocas de atenção permitem explorar o ambiente sequencialmente.

O sistema permite também que sejam extraídas algumas informações sobre a distribuição dos objetos na cena, antes de segmentá-los, através da extração dos momentos centrais da cena global. Conforme trabalho apresentado por Giuliano, (1961) é possível descobrir qual o quadrante, tomando-se o centro de massa da cena como referência, apresenta mais massa ou menos massa. Também é possível calcular a orientação do eixo principal da cena ou seu raio de giração, o que significa, conforme pode ser visto no apêndice 2, a localização mais provável de se encontrar um objeto dentro da imagem. Tais informações podem guiar a segmentação e restringir o processamento para um foco de tamanho menor. Para isso, necessita-se de um processo de controle para guardar as posição dos objetos, evitando que posições já analisadas sejam processadas novamente. A possibilidade de se restringir a busca a uma região selecionada pode reduzir o tempo. Esse tipo de abordagem é interessante também pois permite selecionar regiões de processamento *acima, abaixo, à esquerda ou à direita*.

2 – APRESENTAÇÃO E DISCUSSÃO DOS FUNDAMENTOS DO MODELO

Nesta seção, são descritos os principais fundamentos que guiaram o desenvolvimento desta tese tais como o uso de múltiplos canais de características *bottom-up*, a relação entre eles, o uso de informações *top-down*, bem como sua integração com as informações *bottom-up*. A escolha desses princípios foi inspirada na combinação de evidência sobre o sistema visual e a percepção humana e também analisando-se as estratégias computacionais de modelos anteriores que melhor se adaptam aos objetivos a serem atingidos.

2.1 – Aquisição das imagens

Uma imagem contendo um número de padrões desconhecidos é captada por uma câmera e serve de entrada ao sistema. Os modelos de imagem tratadas computacionalmente são mais simples que as imagens reais pois são discretos, bidimensionais e limitados em extensão e em cores ou níveis de cinza. Segundo Costa, (1996) a imagem real é contínua, possui extensão arbitrária e um número arbitrário de níveis de cinza.

Como suporte para a aquisição das imagens foi utilizado um equipamento específico, mostrado na figura 5.1, composto por uma câmera CCD, um sistema de iluminação, uma “*frame grabber*” para a aquisição da imagem e de um microcomputador. O *software* para analisar o sistema foi desenvolvido em Matlab.

Figura 5.1 – Sistema de aquisição de imagem

Um dos problemas relacionados à obtenção da imagem dos objetos é a quantização espacial devido ao processo de digitalização da imagem, podendo causar perda de definição nas bordas do objeto. Para uma versão do objeto em outra orientação a quantização pode produzir resultados diferentes. A única forma de minimizar o erro de quantização espacial é aumentar o número de elementos de imagem para a mesma área, o que geralmente significa maiores custos.

Neste trabalho, as imagens foram adquiridas com uma resolução radiométrica de 256 níveis de cinza. Os objetos estão distribuídos aleatoriamente e podem aparecer dispostos em diferentes tamanhos ou orientações

A metodologia apresentada neste trabalho foi desenvolvida para operar em cenas simples, contendo um número limitado de objetos distribuídos aleatoriamente e para abranger uma classe de tarefas envolvidas com objetos bidimensionais ou tridimensionais que possam ser projetados em duas dimensões mantendo entretanto suas características significativas. Como exemplos destes últimos, temos objetos cuja altura seja irrelevante quando comparada com a sua largura (por exemplo arruelas, porcas, parafusos, chaves, talheres, pratos, etc.). Para isso foi criado um banco de imagens compostas de objetos geométricos simples como quadrado, círculo, triângulo, elipse, pentágono, etc. e um banco composto de ferramentas (70 imagens). Os objetos também são considerados segmentados, não oclusos, e que conseqüentemente podem ser analisadas utilizando técnicas baseadas em características globais.

A figura 5.2 mostra uma imagem básica de uma cena composta por objetos diferentes. Na cena podem ser vistos sete objetos diferentes: Circuito Integrado (CCI ou CI), Chave Hallen (L ou HLL), Círculo (CIR), Chave Inglesa (CHB), Chave de Cano (CHI), Alicate (ALI) e Sargento (C ou CCC). O “X” definido em cada objeto marca o centro de gravidade (centróide), enquanto que “G” refere-se ao centro de gravidade global da cena.

Figura 5.2 – Cena com objetos.

2.2 – Segmentação da Imagem

Na detecção e reconhecimento de objetos bidimensionais ou tridimensionais por sistemas de Visão por Computador, a imagem de entrada é frequentemente simplificada, gerando-se uma imagem de saída cujos pixels tendem a possuir valores mais altos se eles fazem parte de um objeto de interesse e valores baixos caso contrário (Haralick & Shapiro, 1992). Para realmente reconhecer um objeto, existe primeiro a necessidade de identificar regiões na imagem que possuam potencial para ser parte do objeto.

A segmentação é o processo que divide a imagem em suas partes constituintes ou objetos. A segmentação é um dos elementos mais importantes na análise automatizada de imagens porque é neste passo que objetos e outras entidades de interesse são extraídas da imagem para processamento posterior, tal como descrição e reconhecimento (Gonzales & Wintz, 1987).

A idéia de segmentação teve suas raízes na psicologia Gestalt que estudou as preferências exibidas pelos seres humanos em agrupar ou organizar conjuntos de formas dispostas no campo visual, expressas em termos de características como proximidade, similaridade e continuidade. Em Visão Computacional o agrupamento de partes de uma imagem generalizada em unidades homogêneas que compartilham uma ou mais características resulta em uma imagem segmentada.

É muito difícil dizer ao programa de computador o que constitui uma segmentação “significativa”. Em geral os procedimentos gerais de segmentação tendem a obedecer as seguintes regras (Haralick & Shapiro, 1992), embora na prática seja difícil alcançar todas essas propriedades desejáveis:

- Regiões de segmentação de imagem devem ser uniformes e homogêneas com relação à alguma característica tal como nível de cinza ou textura.
- O interior das regiões deve ser simples e sem muitos buracos pequenos.
- Regiões adjacentes de uma segmentação devem possuir valores adjacentes significativamente diferentes com relação às características nas quais elas são uniformes.
- As bordas de cada segmento devem ser simples, não serrilhada e espacialmente precisa.

O processo de segmentação pode ser determinado de maneira dependente ou independente do contexto, existindo vantagens nesses dois tipos de abordagens. Os algoritmos de segmentação geralmente baseiam-se em uma das duas propriedades básicas de

valores de nível de cinza: descontinuidade e similaridade. Técnicas para detecção e segmentação de bordas podem ser vistas em Ballard & Brown, (1982). As principais abordagens baseiam-se em binarização (*thresholding*), crescimento de região (*region growing*) e estratégias “divide - e - une” (*splitting and merging*). O crescimento de região é uma técnica global que pode ser feito também através da limiarização (“*Thresholding*”).

O algoritmo de segmentação escolhido para este trabalho foi o *Crescimento de Região* que baseia-se na distribuição de sementes (ordenada ou aleatória) pela imagem, através das quais analisam-se e agrupam-se os pixels vizinhos de acordo com uma regra de similaridade, iterativamente. Os métodos iterativos em geral são muito lentos para a maioria das aplicações e novas técnicas ainda estão sendo desenvolvidas. Uma vez que cada grupo de pixels vizinhos é analisado iterativamente, o número de iterações é governado pela parte mais larga do objeto e o tempo envolvido em cada iteração é governado pelo número de pixels existente em cada imagem do objeto. No algoritmo utilizado um nível de *threshold* foi adotado a fim de eliminar sombras e facilitar a computação.

O crescimento de região é um procedimento que agrupa pixels ou sub-regiões em regiões maiores. A forma mais simples desta abordagem consiste em determinar pontos de “semente” e a partir destes pontos, pixels vizinhos que compartilham propriedades similares (por exemplo nível de cinza, cor, textura) são agregados em regiões. A vantagem é que a imagem não precisa ser homogênea, pois as suas características são previamente analisadas e incluídas nos descritores de semelhança. Os problemas imediatos com esta abordagem são a eleição das sementes iniciais que representem adequadamente as regiões de interesse e a seleção de propriedades adequadas para a inclusão dos pontos em várias regiões durante o processo de crescimento. A seleção de um ou mais pontos de semente é dependente da natureza do problema. Em aplicações militares, por exemplo, em imagens em infravermelho, os alvos de interesse são mais quentes e mais brilhantes que o fundo e pontos brilhantes podem ser usados como sementes. A seleção das propriedades que estabelecem os critérios de similaridade depende do tipo de dados disponíveis. Por exemplo as imagens de satélite utilizam a informação de cor.

Após as regiões terem sido definidas, podem ser realizadas uma grande variedade de medidas que constituem rotulagem de atributos ou uma etapa de extração de características.

2.3 – Escolha dos atributos

Os atributos usados para “guiar” a atenção foram escolhidos com base nas evidências provenientes do estudo da percepção e atenção humana (Treisman & Gelade, 1980; Julesz, 1984): tonalidade (média de nível de cinza), tamanho, inclinação (orientação), forma e distância. Todas essas características afetam a nossa percepção e chamam a atenção de maneiras diferentes, dependendo do contexto em que se apresentam em uma cena. A característica distância usada aqui pode fornecer alguma vantagem por exemplo para os objetos mais isolados uns dos outros (Gestalt) embora neste trabalho optou-se por “facilitar” a percepção de objetos próximos ao centro de massa da cena.

Essa escolha das características permitiu definir também um conjunto de atributos que especificam as tarefas a serem realizadas tais como determinar o objeto *maior*, *menor*, *mais perto*, *mais longe*, *mais claro*, *mais escuro*, *muito inclinado*, *pouco inclinado* ou selecionar um objeto específico em função de sua forma. Esses atributos também podem ser combinados em tarefas do tipo “pegar o menor alicate e o mais próximo”, selecionar a chave inglesa mais escura e assim por diante.

Para a extração de todas as características relatadas acima, foi utilizada a teoria dos momentos. Conforme pode ser visto no apêndice 2, através dos momentos podem ser extraídas muitas informações úteis dos objetos tais como posição, orientação, formato e tamanho, mas outras características também podem ser acrescentadas tais como simetria, alongação, etc.

A vantagem dos momentos sobre outras técnicas é que esses descritores de formato tem uma implementação direta e também carregam uma interpretação “física” da forma. Como muitas tarefas necessitam que os objetos possam ser reconhecidos de pontos de vistas arbitrários existe também a necessidade de extração de características que sejam invariantes à escala (contrações ou dilatações), translação (vertical e horizontal) e/ou à orientação (rotação em superfície).

Além de satisfazer as características necessárias para despertar a atenção e realizar tarefas específicas, as características escolhidas também procuraram satisfazer os requisitos necessários para um sistema de visão computacional quanto a:

- 1) Redução da dimensionalidade do padrão de entrada. Se os pixels estiverem totalmente descorrelacionados e possuírem distribuição randômica em todas as imagens de entrada, então características não podem ser vistas, entretanto a distribuição dos pixels que ocorrem freqüentemente podem ser denominadas como características. Através de um conjunto adequado de características, as imagens podem ser descritas reduzindo a dimensão de sua representação. Características são equivalentes à conceitos verbais que permitem processamento simbólico, a menos que essas estruturas abstratas existam nos níveis intermediários da percepção, estruturas complexas não podem ser visualizadas nem compreendidas.
- 2) Permissão de variações e distorções. Através da abstração sobre as distribuições exatas dos pixels, características podem fornecer uma descrição invariante dos objetos.

A característica média de nível de cinza denominada aqui como “Gray” (G) foi calculada para cada objeto (não binarizado) através da seguinte equação:

$$Mc = \frac{1}{m_{00}} \sum_{m=x}^{x2} \sum_{n=y1}^{y2} f(x, y) \quad (1)$$

O tamanho (T) foi extraído em função da área do objeto dada pelo m_{00} (equação 4 do Apêndice 2). A orientação (O) foi calculada em função da equação 6.41 no apêndice 2 e da tabela A2.2. As distâncias (D) foram calculadas em função do centro de massa do objeto e da cena, dado pelas equações 6.5 e 6.6 do Apêndice 2.

Para a escolha da característica de formato (S) foi necessário primeiro realizar um estudo para decidir quais invariantes de momento de Hu ¹ (equações 6.7, apêndice 2) deveriam ser usados. Devido às deformações que ocorrem durante o processo de digitalização e à razão de aspecto da imagem e também devido às transformações realizadas (rotação, translação e escalonamento), os Invariantes de Momento extraídos apresentam uma distribuição, dentro de um valor máximo e mínimo, com uma média e desvio padrão. Para

¹ Devido ao fato dos Invariantes de Hu apresentarem valores muito pequenos e possuir um enorme *range* dinâmico, utilizamos as transformações sugeridas por Master, (1994) para atenuar esses problemas, mostradas na equação 6.14 do apêndice 2

isso levantou-se as curvas de distribuição Gaussiana normalizadas dos valores para cada um dos invariantes, tomando-se como referência 12 cenas² (figuras 5.3 a 5.9).

Figura 5.3 – Distribuição dos valores obtidos para o invariante de Hu “ ϕ_1 ” para 12 cenas.

Figura 5.4 – Distribuição dos valores obtidos para o invariante de Hu “ ϕ_2 ” para 12 cenas.

² Esse número de cenas foi escolhido como referência devido a um estudo anterior com redes neurais treinadas com um algoritmo de *backpropagation*, onde foram escolhidos no treinamento diferentes números de cenas (por exemplo 8, 12, 20, etc..) e diferentes conjuntos de cenas. O número e as cenas que apresentaram melhores resultados na simulação com as outras cenas, foram usadas para levantar as curvas de distribuição (Gonzaga et al., 1998).

Figura 5.5 - Distribuição dos valores obtidos para o invariante de Hu " ϕ_3 " para 12 cenas.

Figura 5.6 - Distribuição dos valores obtidos para o invariante de Hu " ϕ_4 " para 12 cenas.

Figura 5.7 - Distribuição dos valores obtidos para o invariante de Hu " ϕ_5 " para 12 cenas.

Figura 5.8 - Distribuição dos valores obtidos para o invariante de Hu " ϕ_6 " para 12 cenas.

Figura 5.9 - Distribuição dos valores obtidos para o invariante de Hu " ϕ_7 " para 12 cenas.

Através de estudos realizados com estas imagens e também com outros tipos de imagens diferentes, verificou-se que a soma dos invariantes de momentos de Hu apresentam uma melhor distribuição dos valores, conforme mostra a figura 5.10. Isto provavelmente ocorre devido ao fato de que, conforme demonstrado no apêndice 2, os momentos centrais descrevem várias características de formato dos objetos. Como os invariantes de Hu são construídos através da combinação desses momentos centrais, sua soma faz com que várias características de um mesmo objeto sejam recombinaadas dentro de uma mesma característica, favorecendo dessa maneira uma representação única do objeto, ou seja, um *número de formato*.

Figura 5.10 - Distribuição dos valores obtidos para a soma dos invariantes de Hu “φ1+...+φ7” para 12 cenas.

A soma dos invariantes dos objetos definida aqui como φ, e sua respectiva média e desvio padrão são usados então como características representativas dos objetos e serão utilizadas, quando necessário para desempenhar uma tarefa específica, como um modelo de conhecimento prévio (*top-down*) para a forma dos objetos.

Um outro estudo realizado para ver o comportamento dos momentos como características representativas de formato, “interessantes” para guiar a atenção de uma maneira *bottom-up*, foi criar uma matriz de saliência (mapa), não topográfica, construída utilizando-se uma medida de “diferença de formato” entre os objetos da cena, baseado na média da soma dos invariantes de Hu para cada objeto segmentado. A matriz de diferenças pode ser obtida comparando cada vetor de característica de um objeto com o outro, assim:

$$d_{ij}^v = (\overline{\phi_{iv}} - \overline{\phi_{jv}}) \text{ para } v = 1,2,...n \quad (2)$$

onde v representa o índice dos invariantes de momento de Hu.

Portanto, os mapas de saliência individuais são obtidos pela matriz mostrada na equação 3.

$$S_v = \begin{vmatrix} d_{11}^v & d_{12}^v & - & - & - & - & d_{17}^v \\ d_{21}^v & d_{22}^v & - & - & - & - & d_{27}^v \\ | & | & | & | & | & | & | \\ | & | & | & | & | & | & | \\ | & | & | & | & | & | & | \\ | & | & | & | & | & | & | \\ d_{71}^v & d_{72}^v & - & - & - & - & d_{77}^v \end{vmatrix} \quad (3)$$

O mapa de saliência global é obtido pela soma de cada característica individual, para todos os objetos, normalizado pelo valor máximo. Assim, cada elemento da matriz do mapa de saliência global pode ser calculado pela equação (4) abaixo:

$$m_{ij} = \frac{\sum_{v=1}^n d_{ij}^v}{\max(\sum d_{ij})} \quad (4)$$

Essa medida de “conspicuidade” (C_{ij}) pode ser usada para deslocar a atenção para o objeto “mais diferente” através de um mecanismo *bottom-up*.

A figura 5.11 abaixo mostra o mapa de saliência global obtido considerando as diferenças da soma dos invariantes de Hu. Os objetos mais diferentes entre si, considerando os momentos, são a chave hallen e o círculo, pois apresentaram uma maior distância.

Figura 5.11 – Mapa de saliência global (matriz das diferenças) obtido para a média da diferença da soma dos invariantes de Hu para os objetos mostrados na figura 5.2 (Neves et al., 1999).

2.4 – Representação das características em múltiplos mapas

Muitos modelos de atenção visual utilizam a imagem de entrada para inferir diretamente as localizações de interesse (Mozer, 1988). As representações da imagem construídas pelo cérebro humano são muito mais ricas em informação, por isso as imagem de entrada deveriam ser primeiro processadas para se obter um número de mapas. Alguns sistemas reconheceram a importância de se construir mapas de características (Cave & Wolfe, 1990; Ahmad, 1991; Sandon, 1990). Grossberger (1989) apresenta em seu modelo uma estrutura para computar mapas de características *bottom-up* convoluindo a imagem com um banco de filtros orientados. Muitas outras fontes além de orientação podem ser usadas. Em princípio, a imagem de entrada deveria originar muitos canais de processamento, isso poderia garantir que em muitas situações, um elemento interessante da imagem pudesse ser detectado em pelo menos um desses canais. Outros canais que fornecem informações importantes para atenção são cor (mais do que apenas níveis de cinza) e movimento.

Uma maneira de selecionar regiões de interesse baseia-se no modelo de Treisman (1980) e no modelo de Chapman (1990) e Ahmad (1991) que utilizaram mapas de ativação binários. Ao invés de se representar todos os valores possíveis dentro de um único mapa, é possível dividir a faixa de valores entre $[F_{\min}^k, F_{\max}^k]$ em N subintervalos e representar cada valor de $F_{x,y}^k$ fazendo a atividade como em (5):

$$A_{x,y}^{k,m} = \begin{cases} 1 & \text{se } m \cdot \delta \leq F_{x,y}^k < (m+1) \cdot \delta \\ 0 & \text{caso contrário} \end{cases} \quad (5)$$

onde

$$\delta = \frac{F_{\max}^k - F_{\min}^k}{N} \quad (6)$$

Com essas representações binárias é fácil selecionar localizações. Por exemplo é possível detectar a presença de uma característica de interesse ou é possível avaliar a quantidade de localizações que são ajustadas para 1 em um dado mapa de ativação em particular. Existem dois problemas com essa estratégia. O primeiro problema consiste nos esquemas de conexões, tais como “OUs” lógicos globais, que são necessários para computar a atividade total de cada mapa $A^{k,m}$. O outro problema baseia-se na escolha dos

subintervalos, isto é, dos valores de N . Se N for muito grande haverá muitos mapas de características, cada um carregando pouca informação e se N for muito pequeno, muitos valores serão agrupados juntos e não será possível discriminar entre os objetos.

Kock & Ullman (1987) propuseram que uma medida de conspicuidade deveria ser avaliada para cada localização e que a localização de máxima conspicuidade deveria ser selecionada por uma rede WTA (*Winner Take All*) para realizar a troca de atenção.

Sandon (1990) implementou um mecanismo de competição local entre todas as localizações do mapa de característica $F_{x,y}^k$, fornecendo uma medida de conspicuidade $C_{x,y}^k$ que baseia-se no mecanismo “*ON-center, OFF-surround*” que tem o efeito de selecionar regiões no mapa que diferem das regiões vizinhas. Sandon utilizou um operador Laplaciano de escala pequena com o objetivo de enriquecer o contraste e extrair o contorno dos objetos.

Milanese (1993) propôs que esse passo deveria ser utilizado para extrair regiões ao invés de localizações de interesse e sugeriu um operador de escalas maiores, cuja forma e tamanho pudesse se adaptar ao dados a fim de que cada mapa de característica possa atuar como um detector de regiões (*blobs*) orientado.

Um outro aspecto que deve ser considerado em sistemas atencionais refere-se a seleção de uma região ou localização para a próxima troca de atenção e essa tarefa muitas vezes não é trivial devido ao fato de que características provenientes de múltiplos mapas devem ser combinadas para fornecer uma medida global de saliência de uma localização. (Ahmad, 1991) resolveu o problema de uma forma direta, através de uma soma ponderada dos mapas de conspicuidade dada como em (7):

$$A_{x,y} = \sum_{k=1}^K w_k C_{x,y}^k \quad (7)$$

O problema com essa abordagem consiste na escolha apropriada de pesos w_k . Se não se conhece quais características são importantes para distinguir os possíveis alvos dos outros elementos presentes na cena, fazendo $w_1=w_2=\dots=1/K$, a equação acima se reduz a um valor médio dos mapas de características. O problema é que a localização que é conspícua para uma certa característica não é conspícua para outros mapas e o resultado final pode causar um “efeito de média” e fazer com que várias localizações do mapa global possuam o mesmo

valor. Esse efeito causa uma grande perda de informação, que as vezes dificulta a discriminação entre localizações de atividades mais altas.

Milanese (1993) propôs um esquema de integração para os mapas de conspicuidade baseado em um processo de relaxação. Esse processo não-linear não realiza simplesmente uma média sobre os valores locais mas implementa interações competitivas e cooperativas entre diferentes localizações dos mesmos mapas de conspicuidade, bem como entre diferentes mapas. Em paralelo, o processo de relaxação resulta em regiões que contém localizações de alta conspicuidade enriquecidas no mapa global, enquanto que o resto da imagem é gradualmente reduzido. Após algumas poucas interações, o processo de relaxação produz uma imagem binária que claramente identifica as regiões de interesse selecionadas.

Devido a ênfase dada aos componentes *bottom-up* da atenção, a estrutura do modelo apresentado neste trabalho baseia-se na organização de canais *data-driven*. O potencial do sistema está na capacidade de selecionar regiões na imagem baseado na presença de múltiplos canais de processamento. Os objetivos desses canais é fornecer múltiplas estratégias: para todas as imagens de interesse pelo menos um canal deve ser capaz de selecionar características salientes. Por isso, deveriam haver tantos mapas independentes quanto fosse possível, incluindo todas as características conhecidas que são extraídas no córtex visual.

Os mapas de características do sistema apresentado aqui, baseiam-se em características extraídas através da teoria dos momentos. Cada um desses mapas é representado de maneira retinotópica, ou seja, mantendo a mesma disposição dos objetos da imagem original. O valor do mapa de característica na localização (x,y) dada pelo centro de massa do objeto, será assim representada pelo símbolo $F_{x,y}^k$, onde K é o índice que identifica o mapa de característica.

Uma primeira tentativa de representar as características provenientes de momentos em mapas de saliência baseou-se no modelo apresentado por Cave & Wolfe (1990)³ que calcula a probabilidade de cada localização em conter o alvo, dada pela equação 1 do capítulo 4 (página 109). Isso é feito combinando-se uma medida de interesse *bottom-up* (*data-driven*) e *top-down* (*knowledge-driven*). No trabalho apresentado em Neves et al. (1999) implementou-se o modelo da busca guiada para tarefas simples como achar o objeto

³ Ver capítulo 4: Seção 2.3 e Conclusões (item 3)

mais alongado em uma cena semelhante a figura 5.2 e também determinar tarefas do tipo “encontrar o círculo” ou “encontrar os menores objetos”. Os resultados obtidos são mostrados nas figuras 5.12 e 5.13, onde os mapas de saliência global mostram apenas as localizações com maiores atividades.

Apesar das tarefas terem sido realizadas de maneira simples e com sucesso, o grande problema com essa abordagem consiste em se ponderar as características de acordo com a efetividade das mesmas em se detectar o alvo. Também há uma limitação na representação do alvo desejado (componente *top-down*) em função de um valor apenas.

Figura 5.12 – Mapa de saliência global (GSM) baseado no modelo de Cave & Wolfe (1990) para a tarefa “determinar o objeto mais alongado” para uma cena semelhante a figura 5.2 (Neves et al., 1999).

Figura 5.13 – Exemplo de uma imagem (A) e as tarefas (B) determinar o menor círculo, (C) encontrar os círculos, (D) encontrar os menores objetos. Os mapas resultantes mostram os resultados obtidos para os objetos mais semelhantes aos alvos (Neves et al., 1999).

2.5 – Mapas de Características Primários x Intermediários

No sistema atencional os mapas de características definem um nível de chaveamento do estágio pré-atentivo para o estágio atento da percepção. É muito importante definir esse nível na hierarquia de processamento. Existe muito debate na comunidade de Psicologia⁴ porque diferentes experimentos fisiológicos vêm sustentando hipóteses para ambos os estágios (Milanese, 1993).

Os principais modelos psicológicos de atenção tais como a teoria dos *textures* de Julezs e a teoria de integração de características de Treisman & Gelade (1980) são exemplos de modelos de seleção primários (*early selection*). Eles propuseram que o processamento atento seleciona informações processadas por detectores de baixo nível tais como cor e orientação, entretanto outros experimentos sugeriram que o gargalo de processamento, que motivou a necessidade de processamento atento, podem também ser encontrados em um estágio posterior, após a categorização ou associações semânticas.

Estudos realizados por Posner et al (1987) apud Milanese (1993)⁵ relacionaram atenção visual espacial com outras formas de atenção tais como auditiva e lingüística, permitem concluir que todas as áreas do córtex contribuem para o processo de seleção, desde os mapas de orientação de baixo nível, das colunas corticais em V1, os contornos ilusórios extraídos em V2, as cores e formas representadas em V4 e os campos de movimento no Temporal Médio (TM). É provável que a informação de categorização (rotulagem) obtida no Temporal Inferior (TI) permita comparações fáceis entre múltiplos objetos no nível semântico. Treisman (1990) propôs que os atributos semânticos tais como nomes e classes são armazenados em termos de arquivos de objetos.

O objetivo de um sistema que baseia-se em atenção é detectar os contornos ou primitivas das regiões que são mais interessantes. Isso é feito calculando-se atributos para cada primitiva, os quais representam os mapas de características do sistema. Esses atributos podem ser por exemplo:

- *Contornos*: (1) orientação obtida considerando as extremidades de cada contorno, (2) comprimento dos contornos, (3) contraste médio, obtido pela soma de todas as

⁴ Para uma revisão ver Allport (1989) .

⁵ M.I. Posner , A W. Inhoff, F.J. Friedrich & A.Cohen – “Isolating Attentional Systems: A cognitive-Anatomical Analysis” – Psychobiology, V.15, N.2, 1987, pp. 107-121.

respostas locais do detector de borda ao longo do contorno e (4) curvatura média ao longo do contorno.

- *Regiões*: (1) superfície, (2) média de nível de cinza sobre as regiões, (3) perímetro de uma região e (4) alongação de uma região.

O modelo apresentado neste trabalho baseia-se em características intermediárias, utilizando informações de formato de regiões, baseada em momentos e outras informações derivadas como orientação, escala e circularidade, entre outras. Fazendo-se uma analogia com o sistema visual humano, os mapas de características não são mapas descontínuos como os computados em V1, mas representam entidades macroscópicas tais como contornos e regiões, como ocorrem nas áreas extra-estriadas tais como V2. Um conjunto correspondente de mapas de características são obtidos armazenando-se na localização (x,y) definida pelo centro de massa do objeto o k-ésimo atributo da região.

Milanese (1992) utiliza como primitiva contornos suavizados e regiões uniformes. Primeiro, primitivas de contorno são extraídas utilizando um detector de borda, depois é aplicado um procedimento para determinar contornos conectados. Os contornos são representados de uma maneira lista simbólica, facilitando o *splitting and merging*. Uma característica importante que pode ser facilmente computada é a curvatura e expressa uma importante fonte de informação.

A figura 5.14 extraída de Milanese (1993) mostra primitivas de contorno e regiões, obtidas com esse procedimento em uma imagem contendo três objetos principais sobre um fundo não-uniforme, compostos de muitos itens semelhantes. As primitivas da região são extraídas com um algoritmo clássico de crescimento de região (*Region Growing*) onde uma nova região é definida se o valor de um pixel difere de sua vizinhança de um certo *threshold*. Uma representação simbólica tipo grafo pode ser construída também para regiões, o que permite facilmente a avaliação das relações espaciais. Isso pode ser usado para melhorar os resultados do procedimento: se o tamanho da região está abaixo de um *threshold fixo*, ela pode ser fundida com a vizinhança mais similar.

Figura 5.14 – Exemplos de mapas de característica de nível intermediário. (A) - Imagem de entrada. (B) Primitivas de Contorno. (C) primitivas de Regiões (D) mapa de característica “comprimento dos contornos”.

Os tipos de características descritas acima podem ser usadas para extrair as primitivas mais interessantes, definidas como contornos e regiões cujos atributos sejam diferentes das primitivas remanescentes. Uma vantagem desse tipo de abordagem é que a medida de interesse está diretamente relacionada com as primitivas. Uma outra vantagem é que a região de atenção é definida para entidades macroscópicas, que revelam a presença de objetos inteiros, em oposição aos detalhes de cantos e bordas.

A principal desvantagem para esse tipo de abordagem é a falta de plausibilidade biológica. Embora grandes contornos sejam certamente extraídos em algum lugar do córtex, o mecanismo proposto aqui foi apenas motivado em bases computacionais. Esse aspecto é ainda mais crítico para regiões, para quais pouco se conhecesse sobre a implementação biológica. O simples algoritmo de crescimento de região foi escolhido com base na simplicidade de implementação e sua eficiência computacional e não possui certamente uma

equivalência biológica nos sistemas de visão humana⁶, embora apresente plausibilidade em relação a teoria de percepção Gestalt.

Os algoritmos para ambos os tipos de primitivas (contornos e regiões) são também muito sensíveis a vários parâmetros, tais como *thresholds* de nível de cinza. Isso significa que para cada tipo de imagem, devem ser definidos parâmetros diferentes. Finalmente, os resultados obtidos em ambos os algoritmos de extração das primitivas nem sempre fornecem uma segmentação ideal da imagem em objetos físicos.

Apesar dessas limitações, neste trabalho optou-se por trabalhar com regiões e mapas de características intermediários a fim de se avaliar o potencial das características de formato fornecidas pelos momentos como características atencionais capazes de despertar e guiar a atenção. O modelo entretanto pode ser aprimorado para ser aplicado em imagens mais complexas, por exemplo em imagens coloridas ou seguindo um processamento primário de detecção de bordas, agrupamento e definição de regiões (*early processing*).

2.6 – Mapas de Conspicuidade

A fim de selecionar localizações “interessantes” em uma imagem, os mapas de características devem ser processados por algum “operador de interesse” e deve ser extraído um número correspondente de mapas de conspicuidade.

As representações extraídas através dos mapas oferecem múltiplas maneiras de discriminar a informação. Se cada mapa de característica representa uma faixa de valores possíveis para as características em cada região, então o processo de seleção consiste em se determinar diferentes medidas de interesse para os elementos do mapa. Dois aspectos de seleção devem ser considerados:

⁶ Grossberg & Mingola (1989) apresentam uma abordagem mais plausível para a atenção pré-atentiva, em particular para o tipo de processamento que ocorre nas áreas visuais V1, V2 e V4. A idéia básica de sua arquitetura é possuir dois módulos separados: um especializado em computar as bordas dos objetos e outro especializado em detectar o interior de regiões uniformes. Esses dois sistemas são descritos em um único formalismo baseados em ligações cooperativas e competitivas de *feedback*. Os mecanismos simulam os processamentos realizados pelas células simples, complexas e hipercomplexas e essa complexa arquitetura apresenta mecanismos neurais mais plausíveis.

- Grau de distinção entre as características, isto é, quão diferentes elas são dos outros valores dentro do mesmo mapa de característica.
- As relações espaciais de uma localização com relação a outras localizações do mapa.

A importância do primeiro aspecto é direta. No VISIT de Ahmad (1991), por exemplo, as localizações dos mapas de ativação que possuem as menores atividades são utilizadas para identificar as regiões mais distintas. Se ao invés, os mapas de características puderem codificar a faixa de todos os valores possíveis, a distinção de uma localização pode ser calculada comparando seus valores com os valores das outras localizações. Como exemplos dessa abordagem pode-se citar os trabalhos de Cave & Wolfe (1990) e Milanese (1992). Essa abordagem assume um esquema de conexões entre todos os elementos do mapa.

As medidas representadas pelos mapas de características F^k são usadas para detectar localizações de interesse por meio dos mapas de conspicuidade. $C^k_{x,y}$ está relacionado com a diferença entre o valor de $F^k_{x,y}$ e os outros valores $F^k_{u,v}$ dos objetos localizados nas outras posições, assim:

$$C^k_{x,y} = \sum |F^k_{x,y} - F^k_{u,v}| \quad (8)$$

Com $k=1 \dots k$ características.

Os mapas de conspicuidade devem apresentar uma medida de interesse normalizada entre o intervalo [0 1]. Isso pode ser feito com o seguinte procedimento:

$$\bar{C}^k_{x,y} = \frac{C^k_{x,y}}{C_{\max}} \quad (9)$$

O procedimento de normalização descrito acima soluciona o problema de normalização dos mapas de conspicuidade sem alterar as razões entre eles.

A figura 5.15 mostra a relação dos mapas de conspicuidade com a base de dados criada utilizando os conceitos da lógica *fuzzy* que serão detalhados na próxima seção.

Figura 5.15 – Extração de mapas de características e dos mapas de conspicuidade e suas relações com a base de dados construído a partir dos conceitos da lógica *fuzzy*.

2.6.1 - Mapa de conspicuidade de distância

A grande dificuldade de se conseguir uma métrica para medir a conspicuidade de distância é que esta característica não é totalmente independente das outras. A matriz de saliência para a distância é uma medida da distribuição dos objetos dentro da cena e depende dessa configuração. Essa distribuição dos objetos na cena depende diretamente e indiretamente de vários fatores tais como:

- tamanho da cena
- tamanho dos objetos
- número dos objetos presentes na cena
- forma e orientação dos objetos que afetam suas distribuições espaciais
- localização dos objetos (posição) dentro da cena

Todas essas informações são também de natureza espacial e afetam tanto o centro de massa dos objetos como o centro da massa da figura que são utilizadas para medir a conspicuidade.

A fim de que esses parâmetros pudessem ser considerados, optou-se por utilizar a distância entre os objetos, em relação ao centro de massa da cena, e não simplesmente as distâncias entre os objetos. Conforme mostrado anteriormente, a maneira utilizada para calcular as conspicuidades de distância dos objetos é através da matriz das diferenças. Esse

Trabalho de Conclusão de Curso
TCC 2024

cálculo é interessante pois permite analisar para cada objeto, as contribuições das diferenças individuais em relação aos outros objetos, e não apenas a soma dessas contribuições.

As conspicuidades são calculada pela seguinte rotina:

```
for i=1:num
for j=1:num
sald(i,j)=abs(d(i)-d(j));
end
Cd(i)=sum(sald);
end
```

onde num representa o número de objetos presentes na cena, i o índice para os objeto, sald é a matriz de saliência de distância e Cd é a conspicuidade de distância que representa a soma das diferenças de distância de um objeto em relação aos outros.

Devido ao fato de se usar valores absolutos, a matriz resultante é simétrica e suas linhas são iguais às suas colunas. Analisando-se por exemplo a primeira coluna (ou a primeira linha), é possível ver as contribuições das diferenças entre o objeto 1 da cena em relação ao demais e a soma de todas as atividades dentro da coluna, fornece uma medida de conspicuidade total. Através de uma normalização das conspicuidades, através da equação (9) é possível avaliar a contribuição de cada objeto dentro do mapa.

Como um exemplo, consideremos a figura 5.16 a seguir, onde os valores mostrados nos eixos horizontal e vertical correspondem as coordenadas espaciais x e y .

Figura 5.16 – Cena contendo um objeto mais distante em relação ao grupo.

As distâncias para os objetos são:

$D(i) = 58.1400 \quad 59.2282 \quad 59.0346 \quad 48.9616 \quad 59.3123 \quad 53.7954 \quad 89.2792$

A matriz de saliência de distância resultante é mostrada na tabela 5.1.

Tabela 5.1 – Matriz de saliência de distância para a cena mostrada na figura 5.16

SALD(i,j)	1	2	3	4	5	6	7
1	0	1.0882	0.8982	9.1784	1.1723	4.3446	31.1392
2	1.0882	0	0.1936	10.2666	0.0841	5.4329	30.051
3	0.8946	0.1936	0	10.0730	0.2777	5.2393	30.2446
4	9.1784	10.2666	10.073	0	10.3507	4.8338	40.3176
5	1.1723	0.0841	0.2777	10.3507	0	5.517	29.9669
6	4.3446	5.4329	5.2393	4.8338	5.517	0	35.4838
7	31.1392	30.051	30.2446	40.3176	29.9669	35.4838	0
Soma-Cd(i)	47.8174	47.1164	46.9228	85.0202	47.3687	60.8513	197.2031

Conforme resultados apresentados na tabela 5.1, o objeto 7 apresenta uma maior conspicuidade devido ao fato de estar com uma diferença de distância em relação ao centro de massa bem maior, com relação a todos os demais objetos.

Uma outra informação interessante que pode ser analisada a partir da tabela de diferença de distância é conseguida através da adoção de um valor de *threshold* que representa um valor mínimo para que as diferenças sejam consideradas conspícuas. O valor adotado para o *threshold* considerou conspícuas as diferenças individuais dentro do mapa de saliência *salid* (i,j) cujo valor seja maior que 30% do valor máximo medido para as distâncias e não conspícuas as diferenças que apresentam valores abaixo dessa faixa. Dessa forma, cada objeto contribui (valor 1) ou não contribui (valor 0) de acordo com o grau de diferença em relação a um objeto específico. O valor de *threshold* representa um valor de distância mínima que pode ser considerado perceptível. Para essa análise adotou-se como referência 30% do valor máximo de distância, entretanto outros valores podem ser escolhidos de acordo com a precisão desejada. A tabela 5.2 a seguir mostra a matriz de saliência normalizada para a cena mostrada na figura 5.16.

Tabela 5.2 – Matriz de saliência para objetos da figura 5.16, normalizada em relação a $T=0.3d_{max}$

SALD'(i,j)	1	2	3	4	5	6	7
1	0	0	0	0	0	0	1
2	0	0	0	0	0	0	1
3	0	0	0	0	0	0	1
4	0	0	0	0	0	0	1
5	0	0	0	0	0	0	1
6	0	0	0	0	0	0	1
7	1	1	1	1	1	1	0
Soma C'(i)	1	1	1	1	1	1	6

Conforme pode ser visto, o objeto 7 apresentou uma diferença de conspicuidade relevante em relação a todos os outros seis objetos, e por sua vez, cada objeto se diferenciou apenas dele. Esse resultado pode ser interpretado da seguinte maneira: a conspicuidade de um objeto está sendo medida agora pelo número de objetos diferentes que satisfazem um valor de *threshold* mínimo, sendo proporcional a este número. Quanto maior o número de objetos diferentes, maior a conspicuidade. O objetivo de mostrar esta tabela normalizada é ressaltar o tipo de informação que a matriz de saliência de distância carrega. O valor de *threshold* pode ser usado também como um filtro dentro da matriz de saliência, deixando passar apenas a informação mais relevante. Se a conspicuidade for calculada pela equação 9 considerando apenas as contribuições filtradas do objeto (soma das diferenças significativas), as diferenças entre as conspicuidades dos objetos é ainda mais ressaltada, inibindo um pouco a contribuição na atividade dos objetos menos conspicuos.

Um outro exemplo ilustrado na figura 5.17, mostra uma cena onde os objetos apresentam uma configuração simétrica de distribuição dos objetos, sendo o objeto central o que apresenta uma maior conspicuidade. As distâncias calculadas são: $D(i) = 258.8695$ 257.3953 179.2326 8.5889 177.9297 257.0715 258.439.

Figura 5.17 – Cena contendo objetos em diferentes posições.

A tabela 5.3 mostra os valores obtidos para a matriz de saliência e a tabela 5.4 mostra a normalização, considerando o valor de *threshold*.

Tabela 5.3 – matriz de saliência de distância para a cena mostrada na figura 5.17

Salid(i,j)	1	2	3	4	5	6	7
1	0	1.4742	79.6369	250.2806	80.9398	1.7980	0.4302
2	1.4742	0	78.1627	248.8064	79.4656	0.3238	1.0440
3	79.6369	78.1627	0	170.6436	1.3029	77.8389	79.2067
4	250.2806	248.8064	170.6436	0	169.3408	248.4826	249.8503
5	80.9398	79.4656	1.3029	169.3408	0	79.1418	80.5096
6	1.7980	0.3238	77.8389	248.4826	79.1418	0	1.3678
7	0.4302	1.0440	79.2067	249.8503	80.5096	1.3678	0
Soma-C(i)	414.6	409.3	486.8	1337.4	490.7	409	412.4

Tabela 5.4 – Matriz de saliência para figura 5.17 normalizada com relação a T=0.3%

SALD'(i,j)	1	2	3	4	5	6	7
1	0	0	1	1	1	0	0
2	0	0	1	1	1	0	0
3	1	1	0	1	0	1	1
4	1	1	1	0	1	1	1
5	1	1	0	1	0	1	1
6	0	0	1	1	1	0	0
7	0	0	1	1	1	0	0
Soma C'(i)	3	3	5	6	5	3	3

A tabela acima reflete bem a situação da cena pois conforme pode ser visto, o objeto 1 não apresenta uma diferença significativa de distância em relação aos objetos 2, 6 e 7 e o mesmo ocorre entre estes objetos, daí a conspicuidade ser a mesma para eles. O objeto 4 apresenta uma distância bem diferente em relação aos outros seis objetos da cena.

O *threshold* permite uma maior flexibilidade do sistema e afeta a sua sensibilidade. Se for considerado um limiar de 30% do valor da distância máxima para a figura 5.18, os objetos serão considerados aproximadamente equidistantes e não apresentaram atividade (conspicuidade filtrada). O valor de T define quão distante pode estar o objeto mais distante em relação aos demais.

Figura 5.18- Cena contendo objetos aproximadamente equidistantes.

A figura 5.18, ilustra uma situação em que os objetos apresentam a mesma forma, mesma orientação, tamanho e média de nível de cinza e aproximadamente estão dispostos de uma maneira equidistantes em relação ao centro de massa da figura.

As distâncias medidas para os objetos ($i = 1 \dots 7$) em relação ao centro de massa são respectivamente:

D (i) = 53.2572 58.7882 54.3483 51.9412 57.1767 57.2642 56.4271

Esses valores são aproximadamente iguais. As atividades (conspicuidades) C ($i = 1 \dots 7$) medidas são mostradas na tabela 5.5. A tabela 5.6 mostra a conspicuidade resultante considerando o limiar escolhido.

Tabela 5.5 – matriz de saliência de distância para a cena mostrada na figura 5.18.

SALD(i,j)	1	2	3	4	5	6	7
1	0	5.5310	1.0910	1.3161	3.9194	4.0069	3.1698
2	5.5310	0	4.44	6.8471	1.6116	1.5241	2.3611
3	1.0910	4.44	0	2.4071	2.8284	2.9159	2.0788
4	1.3161	6.8471	2.4071	0	5.2355	5.3230	4.4859
5	3.9194	1.6116	2.8284	5.2355	0	0.0875	0.7496
6	4.0069	1.5241	2.9159	5.3230	0.0875	0	0.8371
7	3.1698	2.3611	2.0788	4.4859	0.7496	0.8371	0
Soma C(i)	19.0343	22.3148	15.7612	25.6147	14.4320	14.6945	13.6824

Tabela 5.6 – Matriz de saliência para objetos da figura 5.18, normalizada em relação a $T=0.3d_{max}$

SALD'(i,j)	1	2	3	4	5	6	7
1	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0
Soma C'(i)	0	0	0	0	0	0	0

2.6.2 – Mapa de conspicuidade de média de nível de cinza

O mapa de conspicuidade de média de nível de cinza é calculado de maneira análoga ao mapa de conspicuidade de distância. A figura 5.19 ilustra um exemplo.

Figura 5.19 – Cena contendo um objeto conspícuo com relação à média de nível de cinza.

Os valores obtidos para as médias de nível de cinza para todos os objetos são:
 $G(i) = 16 \ 177 \ 161 \ 189 \ 177 \ 177 \ 177$.

A tabela 5.7 à seguir mostra a matriz de saliência para a média de nível de cinza obtida para os objetos da figura 5.19.

Tabela 5.7 – matriz de saliência de média de nível de cinza para a cena mostrada na figura 5.19.

SaliG(i,j)	1	2	3	4	5	6	7
1	0	161	145	73	161	161	161
2	161	0	16	12	0	0	0
3	145	16	0	28	16	16	16
4	173	12	28	0	12	12	12
5	161	0	16	12	0	0	0
6	161	0	16	12	0	0	0
7	161	0	16	12	0	0	0
SomaCg(i)	962	189	237	249	189	189	189

A tabela 5.7 pode ser também normalizada considerando um valor de *threshold* que significa a diferença mínima de tonalidade (média de nível de cinza) que pode ser percebida.

Apesar da percepção humana apresentar diferentes limiares para a percepção de características de tonalidade como claro e escuro, adotou-se por simplificação um *threshold* T de 16 níveis. Assim, a matriz de saliência normalizada resultante para a figura 5.19 é mostrada na tabela 5.8 a seguir.

Tabela 5.8 – Matriz de saliência de média de nível de cinza normalizada para a cena mostrada na figura 5.19.

SaliG(i,j)'	1	2	3	4	5	6	7
1	0	1	1	1	1	1	1
2	1	0	0	0	0	0	0
3	1	0	0	1	0	0	0
4	1	0	1	0	0	0	0
5	1	0	0	0	0	0	0
6	1	0	0	0	0	0	0
7	1	0	0	0	0	0	0
SomaCg(i)	6	1	2	2	1	1	1

2.6.3 – Mapa de conspicuidade de Formato

O mapa de conspicuidade de formato fornece uma medida de similaridade entre as formas, conforme exemplifica a figura 5.20 onde o objeto mais diferente do grupo apresenta uma maior conspicuidade. No cálculo da conspicuidade de formato, foi utilizada como característica de forma a soma dos invariantes de Hu.

Figura 5.20 – Cena contendo objetos com diferentes formatos.

A tabela 5.9 a seguir mostra a matriz de saliência de formato obtida para os objetos mostrados na figura 5.20.

Tabela 5.9 – Matriz de saliência de formato para a cena mostrada na figura 5.20.

SaliF(i,j)	1	2	3	4	5	6	7
1	0	0	1.4211	0	0	0	5.0584
2	0	0	1.4211	0	0	0	5.0584
3	1.4211	1.4211	0	1.4211	1.4211	1.4211	3.6373
4	0	0	1.4211	0	0	0	5.0584
5	0	0	1.4211	0	0	0	5.0584
6	0	0	1.4211	0	0	0	5.0584
7	5.0584	5.0584	3.6373	5.0584	5.0584	5.0584	0
SomaCf(i)	6.4795	6.4795	10.7426	6.4795	6.4795	6.4795	28.9293

Conforme mostrado na tabela 5.9, o círculo é o objeto com o formato mais diferente na cena, portanto ele atrai mais a atenção do que os outros objetos. De maneira análoga aos outros mapas de conspicuidade também pode ser estabelecido um *threshold* mínimo para que dois objetos sejam similares. Na figura 5.20 o objeto 3 apresenta um formato muito parecido

com as outras estrelas, podendo até ser considerado como tal. Como exemplo adotou-se um limiar de $T=2$ e a matriz de saliência normalizada resultante está mostrada na tabela 5.10.

Tabela 5.10 – Matriz de saliência de formato normalizada para a cena mostrada na figura 5.20.

SaliF(i,j)'	1	2	3	4	5	6	7
1	0	0	0	0	0	0	1
2	0	0	0	0	0	0	1
3	0	0	0	0	0	0	1
4	0	0	0	0	0	0	1
5	0	0	0	0	0	0	1
6	0	0	0	0	0	0	1
7	1	1	1	1	1	1	0
SomaCf(i)'	1	1	1	1	1	1	6

2.6.4 – Mapa de Conspicuidade de orientação

A figura 5.21 a seguir ilustra uma cena contendo objetos em diferentes orientações. As orientações calculadas para os objetos (da esquerda para a direita respectivamente) foram em graus: $tetag = -87.7479 -87.7479 -87.7479 -71.5203 -87.7479 -38.5375 -87.7479$. Para calcular as orientações, utilizou-se A tabela 5.11 mostra a atividade de conspicuidade calculada para cada objeto da cena.

Figura 5.21 – Cena com objetos contendo orientações conspícuas.

Tabela 5.11 – Matriz de saliência de orientação para a cena mostrada na figura 5.21.

SaliO(i,j)	1	2	3	4	5	6	7
1	0	0	0	16.2276	0	49.2105	0
2	0	0	0	16.2276	0	49.2105	0
3	0	0	0	16.2276	0	49.2105	0
4	16.2276	16.2276	16.2276	0	16.2276	32.9828	16.2276
5	0	0	0	16.2276	0	49.2105	0
6	49.2105	49.2105	49.2105	32.9828	49.2105	0	49.2105
7	0	0	0	16.2276	0	49.2105	0
SomaCo(i)	65.4381	65.4381	65.4381	114.1209	65.4381	279.0351	65.4381

Para a orientação também pode ser adotado um *threshold* mínimo a partir do qual as diferenças de orientação serão consideradas. No trabalho adotou-se um limiar de $T=15^{\circ}$, entretanto outros valores poderão ser adotados de acordo com a precisão desejada.

2.6.5 – Mapa de Conspicuidade de tamanho

O mapa de conspicuidade de tamanho tem o objetivo de guiar a atenção para o objeto de tamanho mais diferente dentro da cena, conforme pode ser visto na figura 5.22.

Figura 5.22 – Cena contendo objetos de tamanhos diferentes.

Os tamanhos dos objetos medidos foram tam (i) = 101 556 556 556
2075 556 306.

O mapa de conspicuidade resultante pode ser visto na tabela 5.12.

Tabela 5.12 – Matriz de saliência de tamanho para a cena mostrada na figura 5.22.

SaliT(i,j)	1	2	3	4	5	6	7
1	0	455	455	455	1974	455	205
2	455	0	0	0	1519	0	250
3	455	0	0	0	1519	0	250
4	455	0	0	0	1519	0	250
5	1974	1519	1519	1519	0	1519	1769
6	455	0	0	0	1519	0	250
7	205	250	250	250	1769	250	0
SomaCt(i)	3999	2224	2224	2224	9819	2224	2974

Como já era esperado, dentro do contexto apresentado na cena, o objeto mais conspicuo é o maior, seguido pelo menor objeto. Aqui também pode ser adotado um *threshold* que corresponde a uma diferença de tamanho mínima perceptível.

2.7 – Conceitos *Fuzzy* dentro do modelo proposto

Conforme pode ser visto no apêndice 3, a lógica *fuzzy* está relacionada com a resolução de problemas que utilizam conceitos lingüísticos imprecisos tais como “baixo”, “grande” e “médio”.

A lógica *fuzzy* foi proposta inicialmente por L. Zadeh (Zadeh, 1965) como uma extensão dos conjuntos clássicos e possui como base os “conjuntos nebulosos” (Han et al., 1994). Na lógica *fuzzy*, um certo elemento pode pertencer parcialmente a um conjunto com um grau de inclusão ou um grau de pertinência (*membership degree*) que pode assumir valores entre 0 e 1 (Klir & Yuan, 1995). Quando um elemento possui um grau de pertinência “0” , ele não faz parte do conjunto, enquanto que se o seu grau de pertinência for “1”, isto significa que o elemento está totalmente incluído no conjunto. Se os graus de pertinência, entretanto, apresentam valores entre “0” e “1” , isto indica o grau que um elemento pertence

a um conjunto. O grau de inclusão de um conjunto *fuzzy* são determinados pelas chamadas funções de pertinência (*membership functions*). A lógica *fuzzy* será eficaz se todos os passos forem realizados de maneira eficiente, desde a coleção de dados para a determinação dos conjuntos *fuzzy* (*fuzzy sets*) até a interpretação dos resultados provenientes do processo de “desfuzificação”.

Os principais passos envolvidos em um processo *fuzzy* são:

- A determinação dos conjuntos nebulosos
- A determinação das funções de transferência
- “Fuzificação”
- Determinação das operações lógicas *fuzzy*
- Determinação das regras de inferência
- Extração do conhecimento
- “Desfuzificação”

Os conceitos da lógica *fuzzy* serão aplicados neste trabalho com o intuito de fornecer uma metodologia para processar e integrar informações *bottom-up* e *top-down* através de uma rede *fuzzy*, composta de duas camadas. A figura 5.23 a seguir mostra um esquema geral para o modelo proposto.

A primeira camada é composta de dois subsistemas *fuzzy*. O primeiro subsistema *fuzzy bottom-up* tem o objetivo de combinar as diversas informações obtidas pelos mapas de conspicuidade e inferir através de uma base de regras, alguns princípios de percepção visual tais como os definidos pelas leis Gestalt. O sentido dado para Gestalt para estabelecer as regras consiste em se determinar a influência ou contexto global que a presença de uma determinada característica influencia nas outras características. O segundo subsistema *top-down* combina diferentes características de acordo com a relevância das mesmas na realização de diferentes tarefas. Finalmente o subsistema da segunda camada integra a informação dos sistemas anteriores e fornece um índice de saliência (atenção) global.

Figura 5.23 – Visão Geral do modelo proposto, composto de uma rede *fuzzy* de duas camadas. A primeira camada é composta de dois subsistemas *fuzzy*: *bottom-up* e *top-down*, responsáveis por fornecer um índice de atenção dirigido pelas características e pelas tarefas respectivamente.

2.7.1 – Os conjuntos *Fuzzy*

Os conjuntos *fuzzy* de entrada foram escolhidos devido a sua influência em caracterizar os processos de percepção e atenção (conforme mostrado na seção 2.3 deste capítulo) e podem ser vistos na tabela 5.13 e 5.14.

Tabela 5.13 – Conjuntos *Fuzzy* da primeira camada da rede

Subsistema <i>Fuzzy</i>	Conjuntos <i>Fuzzy</i> de entrada	Conjuntos <i>Fuzzy</i> de saída
<i>Bottom-up</i>	<i>Conspicuidade de média de nível de cinza</i>	<i>Saliência Bottom-up</i>
	<i>Conspicuidade de Tamanho</i>	
	<i>Conspicuidade de Orientação</i>	
	<i>Conspicuidade de Formato</i>	
	<i>Conspicuidade de Distância</i>	
<i>Top-down</i>	<i>Média de nível de cinza</i>	<i>Saliência Top-down</i>
	<i>Tamanho</i>	
	<i>Orientação</i>	
	<i>Formato</i>	
	<i>Distância</i>	

Tabela 5.14 – Conjuntos *Fuzzy* para a segunda camada da rede

Subsistema <i>Fuzzy</i>	Conjuntos <i>Fuzzy</i> de Entrada	Conjuntos <i>Fuzzy</i> de Saída
<i>Integração</i>	<i>Saliência Top-down</i>	<i>Saliência Global</i>
	<i>Saliência Bottom-up</i>	

2.7.2 – Extração do Conhecimento

A teoria dos conjuntos *fuzzy* é baseada na afirmação de que os conjuntos clássicos não são naturais, apropriados ou úteis em descrever o comportamento humano e o desenvolvimento de sistemas que atuam como seres humanos foi um dos desafios originais que inspiraram as pesquisas teóricas nessa área. Um outro tipo de motivação foi a criação de programas que pudessem responder apropriadamente ao mesmo tipo de instruções dadas aos seres humanos por outros instrutores para a realização de tarefas específicas. A lógica *fuzzy* permitiu conceitualizar apropriadamente certos aspectos do comportamento humano e são relevantes e úteis para explicar muitos resultados psicológicos (Kochen, 1975).

A base do conhecimento consiste de uma base de dados (funções) e uma base de regras, onde está contido o conhecimento sobre o domínio da aplicação e os objetivos a serem atingidos, de maneira a caracterizar a estratégia do processo e suas metas. Na base de dados ficam armazenadas as definições sobre discretizações e normalizações dos universos de discurso e as definições das funções de pertinência dos termos nebulosos. Para abstrair estes conhecimentos pode-se recorrer a manuais, experimentos ou questionar um especialista.

Neste trabalho, foi elaborado uma pesquisa psicofísica com o objetivo de determinar a função mais adequada para a representação das funções de pertinência de cada conjunto *fuzzy*. Para determinar cada uma das funções de pertinência fez-se um grupo de perguntas para saber a frequência ou peso das respostas.

Conforme ressaltado em Portezani (1999) as sensações estão diretamente relacionadas com os estímulos físicos, os quais são formas de energia, transmitidos aos sensores fisiológicos. Como exemplos podem ser citadas qualidades como azul, duro, quente, etc. Já a percepção está relacionada com o processo psicológico da compreensão de uma sensação, envolvendo um certo padrão de julgamento, onde utilizam-se processos de organização e integração dos estímulos físicos produzidos por sensores fisiológicos. Portezani define ainda Psicofísica como a ciência que procura quantificar a relação entre a variação de certas características dos estímulos físicos (de dimensões físicas) e as respectivas mudanças ocorridas no lado subjetivo do experimento (de características psicológicas). Os métodos utilizados para descrever esta relação são tratados como *métodos psicofísicos*. Devido a natureza subjetiva da percepção e tomada de decisão, a lógica *fuzzy* é adequada

para expressar bem essa subjetividade humana. Alguns trabalhos nesse sentido podem ser vistos em Kochen (1975) e Kokawa et al. (1975).

2.7.2.1 – Determinação das funções de pertinência do subsistema *bottom-up*

Para a realização da pesquisa, foram entrevistadas cinquenta pessoas dentre elas especialistas da área de visão e processamento de imagens incluindo outros profissionais com formação superior em psicologia, comunicação visual, propaganda e marketing, fotografia, etc.

O objetivo desses testes psicofísicos, feito a partir de um questionário, foi extrair o conhecimento humano em questões subjetivas que expressem qualidades de:

- 1) Conspicuidade de tamanho
- 2) Conspicuidade de média de nível de cinza
- 3) Conspicuidade de distância
- 4) Conspicuidade de orientação

Os resultados obtidos com esse teste serviram para a formação da base de funções dos conjuntos *fuzzy* “Conspicuidade de tamanho”, “Conspicuidade de Cinza”, “Conspicuidade de distância” e “Conspicuidade de orientação”, no subsistema *bottom-up*. As funções de pertinência determinam o grau de pertinência de uma determinada variável dentro do correspondente conjunto *fuzzy* e são usadas para modelar o comportamento da percepção humana, mas outros tipos de pesquisas podem ser realizadas no modelamento das funções, de acordo com a aplicação e objetivo do sistema.

2.7.2.1.1 – Determinação das Funções de pertinência para o conjunto *fuzzy* conspicuidade de tamanho

Quando calcula-se a conspicuidade de um objeto em relação aos outros, é possível detectar se uma determinada característica é muito diferente em relação às outras. Por exemplo, se no mapa de característica “tamanho” um objeto apresenta um valor de conspicuidade *muito pequeno*, isto significa que seu tamanho difere pouco dos demais objetos presentes na cena. Por sua vez, se a conspicuidade é *alta*, o objeto em questão apresenta um tamanho *muito* diferente dos demais (ou muito menor ou muito maior).

Para se determinar as funções de pertinência do conjunto *fuzzy* conspicuidade de tamanho, foi elaborado um teste psicofísico simples, o qual consistiu em mostrar para um

grupo de pessoas uma imagem análoga à figura 5.24, contendo objetos de diferentes tamanhos. Os tamanhos dos objetos foram normalizados (de 1 a 10, do menor para o maior objeto) e as pessoas foram requisitadas a especificar em relação ao primeiro objeto qual (ou quais objetos) é pouco conspícuo (diferente), médio conspícuo ou muito conspícuo. Cada teste foi feito de maneira independente e alternada em relação aos outros tipos de testes para avaliar as conspiciuidades de orientação, cinza e distância. As conspiciuidades também estão normalizadas no intervalo [0 1], sendo 0 para o objeto mais semelhante ao primeiro objeto e 1 para o objeto mais diferente (conspícuo) considerando uma característica específica.

Figura 5.24 – Figura análoga à utilizada no teste psicofísico para medida do grau de atenção em função da diferença de tamanho dos objetos, com relação ao primeiro objeto.

Através de uma análise das respostas obtidas é possível determinar o grau de conspiciuidade em função dos tamanhos e das conspiciuidades normalizados, conforme pode ser visto nas tabelas 5.15, 5.16 e 5.17. As figuras 5.25, 5.26 e 5.27 mostra um gráfico das frequências das respostas em função dos tamanhos normalizados (1 a 10).

Tabela 5.15 – Respostas obtidas para a característica “Pouca conspiciuidade” de tamanho

Tamanhos Normalizados	Pouca Conspiciuidade (diferença) % Votos
1	34 %
2	34%
3	29%
4	3%
5	0%
6	0%
7	0%
8	0%
9	0%
10	0%

Figura 5.25 – Gráfico resultante da avaliação do questionário para o termo “pouca conspicuidade” de tamanho.

Tabela 5.16 – Respostas obtidas para a característica “Média conspicuidade” de tamanho

Tamanhos Normalizados	Média Conspicuidade (diferença) % Votos
1	0 %
2	0 %
3	20 %
4	32 %
5	32 %
6	12 %
7	4 %
8	0%
9	0%
10	0%

Figura 5.26 – Gráfico resultante da avaliação do questionário para o termo “Média conspicuidade” de tamanho.

Tabela 5.17 – Respostas obtidas para a característica “Muita conspicuidade” de tamanho

Tamanhos Normalizados	Muita Conspicuidade (diferença)
	% Votos
1	0 %
2	0 %
3	0 %
4	3,5 %
5	5,4 %
6	12 %
7	17,5 %
8	19,3 %
9	19,3 %
10	23%

Figura 5.27 – Gráfico resultante da avaliação do questionário para o termo “Muita conspicuidade” de tamanho.

A figura 5.28 mostra a representação gráfica para as funções de pertinência do conjunto *fuzzy* “conspicuidade de tamanho” em função dos tamanhos normalizados no intervalo [1 10].

Figura 5.28 – Funções de Pertinência para os termos “Pouca”, “Média” e “Muita” do conjunto *fuzzy* de entrada “Conspicuidade de Tamanho” em função dos tamanhos normalizados.

Um outra normalização é necessária para transformar a escala em função do tamanho para uma escala em função da conspicuidade normalizada dentro do intervalo [0 1]. Para isso utilizou-se a transformação em escala onde a conspicuidade do objeto 1 em relação a ele mesmo é mínima (0) e a conspicuidade do objeto 1 em relação ao objeto 10 é máxima (1). A

figura 5.29 mostra as funções de pertinência para a conspicuidade de tamanho em função da escala transformada de conspicuidade.

Figura 5.29 – Funções de Pertinência para os termos “Pouca”, “Média” e “Muita” do conjunto *fuzzy* de entrada “Conspicuidade de Tamanho” em função da conspicuidade normalizada no intervalo [0 1].

Na figura 5.30 é mostrada uma aproximação para as funções de pertinências da conspicuidade de tamanho.

Figura 5.30 – Aproximação das funções de pertinência do conjunto *fuzzy* “conspicuidade de tamanho”.

2.7.2.1.2 – Determinação das Funções de pertinência para o conjunto fuzzy “Conspicuidade de cinza”

Para se determinar as funções de pertinência para o conjunto fuzzy “conspicuidade de cinza” realizou-se um teste utilizando uma figura análoga à figura 5.31. Durante a realização dos testes as pessoas investigadas foram requisitadas a determinar qual ou quais objetos apresentam pouca, média e muita conspicuidade (diferença) de média de nível de cinza, em relação ao primeiro objeto. Os testes realizados foram realizados de maneira independente e alternados com os outros testes de conspicuidade. Os níveis de cinza escolhidos para a realização do teste foram 15, 47, 79, 95, 159, 191 e 223 e correspondem respectivamente aos índices 1,2...8 relativos a cada tom (do mais escuro para o mais claro).

Figura 5.31 - Figura análoga à utilizada no teste psicofísico para medida do grau de atenção em função da diferença de tonalidade dos objetos em relação ao primeiro objeto.

Através de uma análise das respostas obtidas na pesquisa realizada foi possível determinar o grau de conspicuidade em função das tonalidades e das conspicuidades normalizadas, conforme mostram as tabelas 5.18, 5.19 e 5.20. As figuras 5.32, 5.33 e 5.34 mostram um gráfico das freqüências das respostas em função das tonalidades normalizadas (1 a 8).

Tabela 5.18 – Respostas obtidas para a característica “pouca conspicuidade” de Nível de Cinza

Tonalidades Normalizadas	Pouca Conspicuidade (diferença) - %votos
1	37%
2	37%
3	17%
4	9%
5	0%
6	0%
7	0%
8	0%

Figura 5.32 - Gráfico resultante da avaliação do questionário para o termo “Pouca conspicuidade” de nível de cinza.

Tabela 5.19 – Respostas obtidas para a característica “média conspicuidade” de Nível de Cinza

Tonalidades Normalizadas	Média Conspicuidade (diferença) - %votos
1	0 %
2	3,6 %
3	21,4 %
4	35,7 %
5	35,7 %
6	3,6 %
7	0 %
8	0 %

Figura 5.33 - Gráfico resultante da avaliação do questionário para o termo “Média conspicuidade” de nível de cinza.

Tabela 5.20 – Respostas obtidas para a característica “Muita conspicuidade” de Nível de Cinza

Tonalidades Normalizadas	Muita Conspicuidade (diferença) - %votos
1	0 %
2	0 %
3	0 %
4	3 %
5	6 %
6	20 %
7	32 %
8	39 %

Figura 5.34 - Gráfico resultante da avaliação do questionário para o termo “Muita conspicuidade” de nível de cinza.

A figura 5.35 a seguir mostra a representação gráfica para as funções de pertinência do conjunto *fuzzy* “conspicuidade de nível de cinza” em função das tonalidades normalizadas no intervalo [1 8]. A figura 5.36 mostra as funções de pertinência em função da conspicuidade normalizada no intervalo entre [0 1] .

Figura 5.35 – Representação das funções de pertinência para os termos “Pouca”, “Média” e “Muita” conspicuidade do conjunto *fuzzy* “Conspicuidade de Nível de Cinza” em função dos níveis de cinza normalizados.

Figura 5.36 - Representação das funções de pertinência para os termos “Pouca”, “Média” e “Muita” conspicuidade do conjunto *fuzzy* “Conspicuidade de Nível de Cinza” em função da conspicuidade normalizada entre [0 1].

A figura 5.37 mostra a aproximação realizada para as funções de pertinência do conjunto *fuzzy* de entrada “conspicuidade de nível de cinza”.

Figura 5.37 – Aproximação das funções de pertinência dos termos “Pouca”, “Média” e “Muita” do conjunto fuzzy de entrada “conspicuidade de Nível de Cinza” .

2.7.2.1.3 – Determinação das Funções de pertinência para o conjunto fuzzy “Conspicuidade de orientação”

Para se determinar as funções de pertinência para o conjunto fuzzy “conspicuidade de orientação” realizou-se um teste utilizando a figura 5.38.

Figura 5.38 - Figura análoga à utilizada no teste psicofísico para medida do grau de conspicuidade em função da inclinação dos objetos, em relação ao primeiro objeto.

No teste realizado, as pessoas foram requisitadas a indicarem com relação ao primeiro objeto (da esquerda para a direita), qual ou quais objetos apresentavam pouca, média e muita conspicuidade (diferença) de orientação. Com base nas respostas obtidas, foi possível determinar as funções de pertinências para os respectivos atributos lingüísticos “pouca”, “média” e “muita” conspicuidade do conjunto “conspicuidade de orientação”.

Através de uma análise das respostas obtidas é possível determinar o grau de conspicuidade em função das orientações e das conspicuidades normalizadas, conforme pode ser visto na tabela 5.21. As figuras 5.39, 5.40 e 5.41 mostram um gráfico das frequências das respostas em função das orientações normalizadas (1 a 7).

Tabela 5.21 – Respostas obtidas para “pouca”, “média” e “muita” conspicuidade de orientação

Orientações Normalizadas	Pouca Conspicuidade de Orientação	Média Conspicuidade de Orientação	Muita Conspicuidade de Orientação
1	38%	0%	0%
2	38%	4%	0%
3	24%	36%	0%
4	0%	44%	0%
5	0%	16%	22%
6	0%	0%	39%
7	0%	0%	39%

Figura 5.39 – Gráfico resultante da pesquisa realizada para o termo “pouca” conspicuidade de orientação.

Figura 5.40 – Gráfico resultante da pesquisa realizada para o termo “média” conspicuidade de orientação.

Figura 5.41 – Gráfico resultante da pesquisa realizada para o termo “muita” conspicuidade de orientação.

A figura 5.42 mostra a representação gráfica para as funções de pertinência do conjunto *fuzzy* “conspicuidade de orientação” em função das orientações normalizadas no intervalo [1 7]. A figura 5.43 mostra as funções de pertinência para os termos “pouca”, “Média” e “Muita” conspicuidade de orientação, em função da conspicuidade normalizada no intervalo entre [0 1] . Na figura 5.44 pode ser vista uma aproximação para essas funções de pertinência.

Figura 5.42 - Representação das funções de pertinência para os termos “Pouca”, “Média” e “Muita” conspicuidade do conjunto *fuzzy* “Conspicuidade de orientação” em função das orientações normalizadas.

Figura 5.43 - Representação das funções de pertinência para os termos “Pouca”, “Média” e “Muita” conspicuidade do conjunto *fuzzy* “Conspicuidade de Orientação” em função da conspicuidade normalizada entre [0 1].

Figura 5.44 - Aproximação das funções de pertinência para os termos “Pouca”, “Média” e “Muita” do conjunto *fuzzy* de entrada “conspicuidade de orientação” .

2.7.2.1.4 – Determinação das Funções de pertinência para o conjunto *fuzzy* “Conspicuidade de distância”

Para se determinar as funções de pertinência para o conjunto *fuzzy* “conspicuidade de distância” realizou-se um teste utilizando uma figura análoga à figura 5.45.

Figura 5.45 - Figura análoga à utilizada no teste psicofísico para medida do grau de conspicuidade em função da distância dos objetos em relação ao primeiro objeto.

No teste realizado, as pessoas foram requisitadas a indicarem com relação ao primeiro objeto, qual ou quais objetos apresentavam pouca, média e muita conspicuidade de distância. Com base nas respostas obtidas, foi possível determinar as funções de pertinências para os respectivos atributos lingüísticos “pouca”, “média” e “muita” conspicuidade do conjunto “conspicuidade de distância”.

Através de uma análise das respostas obtidas é possível determinar o grau de conspicuidade em função das distâncias e das conspicuidades normalizadas, conforme pode ser visto na tabela 5.22. As figuras 5.46, 5.47 e 5.48 mostram um gráfico das frequências das respostas em função das distâncias normalizadas (1 a 10).

Tabela 5.22 – Respostas obtidas para “pouca”, “média” e “muita” conspicuidade de distância

Distâncias Normalizadas	Pouca Conspicuidade de Distância	Média Conspicuidade de Distância	Muita Conspicuidade de Distância
1	32,5%	0%	0%
2	32,5%	0%	0%
3	20%	7%	0%
4	12,5%	29%	0%
5	2,5%	39%	2%
6	0%	18%	4%
7	0%	7%	19%
8	0%	0%	25%
9	0%	0%	25%
10	0%	0%	35%

Figura 5.46 – Gráfico resultante da pesquisa realizada para o termo “pouca” conspicuidade de distância.

Figura 5.47 – Gráfico resultante da pesquisa realizada para o termo “Média” conspicuidade de distância.

Figura 5.48 – Gráfico resultante da pesquisa realizada para o termo “Muita” conspicuidade de distância.

A figura 5.49 mostra a representação gráfica para as funções de pertinência do conjunto *fuzzy* “conspicuidade de distância” em função das distâncias normalizadas no intervalo [1 10]. A figura 5.50 mostra as funções de pertinência para os termos “pouca”, “Média” e “Muita” conspicuidade de distância, em função da conspicuidade normalizada no intervalo entre [0 1]. Na figura 5.51 pode ser vista uma aproximação para essas funções de pertinência.

Figura 5.49 - Representação das funções de pertinência para os termos “Pouca”, “Média” e “Muita” conspicuidade do conjunto *fuzzy* “Conspicuidade de tamanho” em função das distâncias normalizadas.

Figura 5.50 - Representação das funções de pertinência para os temos “Pouca”, “Média” e “Muita” conspicuidade do conjunto *fuzzy* “Conspicuidade de Distância” em função das conspicuidades normalizadas no intervalo [0 1].

Figura 5.51 - Aproximação das funções de pertinência dos termos “Pouca”, “Média” e “Muita” do conjunto *fuzzy* “Conspicuidade de Distância” .

2.7.2.1.5 – Determinação das Funções de pertinência para o conjunto *fuzzy* “Conspicuidade de Formato”

Para a determinação das funções de pertinência não utilizou-se um teste específico, embora tal procedimento possa ser realizado para a extração de diversos tipos de conhecimento a respeito das diferenças entre as formas. Por exemplo um determinado formato de um objeto pode chamar mais a atenção do que outros objeto iguais a eles, mas com alguma irregularidade em seu formato. Um estudo mais profundo é necessário para melhor estabelecer os limites de tolerância dentro de uma mesma classe e também entre classes. Muitas vezes pode ser necessário que alguma forma específica tenha prioridade em chamar a atenção com relação às outras presentes na cena. Esses aspectos, entre outros, podem ser também considerados.

Na determinação das funções de pertinência para conspicuidade de formato adotou-se as funções mostradas na figura 5.52 .

Figura 5.52 – Função de Pertinência adotadas para o conjunto *fuzzy* “conspicuidade de formato”.

2.7.2.2 – Determinação das Regras de Inferência *fuzzy* para os conjuntos *fuzzy* de “Conspicuidade ”

Devido ao fato da saliência *bottom-up* ser muito influenciada pelo contexto global dos objetos em uma cena (Gestalt), alguns objetos podem atrair a atenção de uma maneira diferente se os objetos a sua volta mudarem. Por exemplo, na figura 5.53 (c) se um objeto possui um valor de média de nível de cinza muito diferente em relação aos outros, características tais como forma e tamanho podem não ser tão relevantes. Pode-se perceber também na figura 5.53 a influência das distâncias entre os objetos em atrair a atenção.

Através das regras de inferência dentro do subsistema *bottom-up*, é possível relacionar os diferentes graus de pertinência provenientes das cinco características independentes (“conspicuidade de nível de cinza”, “conspicuidade de tamanho”, “conspicuidade de orientação”, “conspicuidade de formato” e “conspicuidade de distância”). Características diferentes apresentam graus diferentes de atenção quando combinadas.

Neste trabalho o subsistema *bottom-up* foi elaborado com 243 regras resultantes da combinação das variáveis linguísticas “pouca”, “média” e “muita” dos cinco conjuntos *fuzzy* de entrada.

Figura 5.53 – Exemplo mostrando como a configuração global de uma cena pode ser dependente do contexto das características presentes na mesma. Em (a) o quadrado atrai a atenção para ele pois possui uma forma mais diferente em relação aos outros objetos. O mesmo ocorre em (b) para a reta com orientação diferente. Em (c), a diferença de nível de cinza é a característica que possui uma maior influência em atrair atenção para um objeto. O objeto escuro neste caso é o que atrai mais atenção. Em (d) a característica diferença de tamanho possui uma influência maior do que a diferença de orientação. Nesse caso, o quadrado com a orientação diferente também atrai atenção para ele porém com menor intensidade.

Para o estabelecimento das regras de inferência que relacionam as diferentes conspicuidades das diferentes características, é necessário realizar um estudo baseado na percepção (observação), a fim de averiguar o relacionamento das variáveis linguísticas de entrada com a variável de saída, que neste caso é um índice de saliência (atenção) *bottom-up*. Para isso foi realizada uma pesquisa empírica com algumas pessoas para saber, quais características, quando combinadas entre si, apresentariam maior relevância (peso) ou seja despertariam um maior grau de atenção *bottom-up*. Através dessa pesquisa foi possível verificar que as características de conspicuidade de tamanho, distância e nível de cinza quando presentes se sobressaem em relação as demais, fazendo com que o nível de atenção para os objetos que possuem essas características sejam mais altos. Com base nas respostas obtidas foi possível montar as regras de inferência *fuzzy*. No estabelecimento das regras, além de considerar a contribuição de cada característica em chamar a atenção (pouca, média

ou muita), adotou-se ainda um peso para cada mapa: $p_1=2$, $p_2=1.5$, $p_3=1$, $p_4=1.2$ e $p_5=1.8$ respectivamente para os mapas conspicuidade de nível de cinza, conspicuidade de tamanho, conspicuidade de orientação, conspicuidade de formato e conspicuidade de distância.

As pesquisas realizadas procurou aproximar a percepção e atenção humana fazendo com o que o sistema reproduzisse uma reação próxima à reação humana, embora um estudo mais detalhado seja necessário para um modelamento mais preciso. A mesma metodologia pode ser aplicada para modelar diferentes comportamentos dependendo da aplicação, por exemplo em outros sistemas a característica orientação pode apresentar uma maior relevância (peso).

Como exemplo de regras, considere a situação do quadrado da figura 5.53 (a) que apresenta a mesma inclinação, o mesmo tamanho e o mesmo nível de cinza. A regras de inferência que se ajustariam à sua descrição dentro do contexto daquela cena específica poderiam ser:

- SE objeto apresenta “*pouca* diferença (conspicuidade) de nível de cinza” E “*pouca* conspicuidade de tamanho” E *pouca* “conspicuidade de orientação” E *média* conspicuidade de forma e *muita* conspicuidade de distância ENTÃO a atenção é “*média*”.
- SE objeto apresenta *muita* conspicuidade de nível de cinza” E “*média*” conspicuidade de tamanho” E *pouca* “conspicuidade de orientação” E *média* conspicuidade de forma e *muita* conspicuidade de distância ENTÃO a atenção é “*média- muita*”.

A figura 5.54 ilustra como as regras de inferência foram obtidas:

Figura 5.54 – Regras de inferência especificadas por um domínio especialista e por literatura relacionada.

Para representar o conjunto *fuzzy* de saída “saliência *bottom-up*” foram escolhidos os seguintes termos lingüísticos: pouca (p), média-pouca (mp), média (m), média-muita (mg) e muita (g).

Na figura 5.55 podem ser vistas as funções de pertinência do conjunto *fuzzy* de saída que determina o índice de saliência *bottom-up* após terem sido combinadas as informações dos cinco conjuntos *fuzzy* de entrada.

Figura 5.55 – funções de pertinência para as variáveis “pouca” (p), “média-pouca” (mp), “média” (m), “média-muita” (mg) e “muita” (g) do conjunto *fuzzy* de saída “saliência *bottom-up*”.

O índice de saliência *bottom-up* (*isb*) é obtido através de um processo de “desfuzificação” que agrega as saídas obtidas para os conjuntos *fuzzy* de entrada, através do método da centróide, apresentado no apêndice 3.

A tabela 5.23 mostra como foram construídas as 243 regras de inferência utilizadas para determinar o índice de saliência *bottom-up* (*isb*).

As figuras 5.56 a 5.60 são exemplos de algumas cenas onde foram calculados os índices de saliência *bottom-up* para os objetos.

Tabela 5.23 – Tabela de regras de inferência para determinação do índice de saliência *bottom-up*

Regras	M Nível de Cinza	Tamanho	Orientação	Forma	Distância	Saliência Bottom-up
1	Pouca	pouca	Pouca	pouca	Pouca	pouca
2	Pouca	pouca	Pouca	pouca	Média	pouca
3	Pouca	pouca	Pouca	pouca	muita	média-pouca
4	Pouca	pouca	Pouca	média	pouca	pouca
5	Pouca	pouca	Pouca	média	média	média-pouca
6	Pouca	pouca	Pouca	média	muita	média-pouca
7	Pouca	pouca	Pouca	muita	pouca	pouca
8	Pouca	pouca	Pouca	muita	média	média-pouca
9	Pouca	pouca	Pouca	muita	muita	média
10	Pouca	pouca	Média	pouca	pouca	pouca
11	Pouca	pouca	Média	pouca	média	pouca
12	Pouca	pouca	Média	pouca	muita	média-pouca
13	pouca	pouca	Média	média	pouca	pouca
14	pouca	pouca	Média	média	média	média-pouca
15	pouca	pouca	Média	média	muita	média-pouca
16	pouca	pouca	Média	muita	pouca	média-pouca
17	pouca	pouca	Média	muita	média	média-pouca
18	pouca	pouca	Média	muita	muita	média
19	pouca	pouca	Muita	pouca	pouca	pouca
20	pouca	pouca	muita	pouca	média	média-pouca
21	pouca	pouca	muita	pouca	muita	média-pouca
22	pouca	pouca	muita	média	pouca	pouca
23	pouca	pouca	muita	média	média	média-pouca
24	pouca	pouca	muita	média	muita	média
25	pouca	pouca	muita	muita	pouca	média-pouca
26	pouca	pouca	muita	muita	média	média
27	pouca	pouca	muita	muita	muita	média
28	pouca	média	pouca	pouca	pouca	pouca
29	pouca	média	pouca	pouca	média	média-pouca
30	pouca	média	pouca	pouca	muita	média-pouca
31	pouca	média	pouca	média	pouca	pouca
32	pouca	média	pouca	média	média	média-pouca
33	pouca	média	pouca	média	muita	média
34	pouca	média	pouca	muita	pouca	média-pouca
35	pouca	média	pouca	muita	média	média-pouca
36	pouca	média	pouca	muita	muita	média
37	pouca	média	média	pouca	pouca	pouca
38	pouca	média	média	pouca	média	média-pouca
39	pouca	média	média	pouca	muita	média
40	pouca	média	média	média	pouca	média-pouca
41	pouca	média	média	média	média	média-pouca
42	pouca	média	média	média	muita	média
43	pouca	média	média	muita	pouca	média-pouca
44	pouca	média	média	muita	média	média
45	pouca	média	média	muita	muita	média
46	pouca	média	muita	pouca	pouca	média-pouca
47	pouca	média	muita	pouca	média	média-pouca
48	pouca	média	muita	pouca	muita	média
49	pouca	média	muita	média	pouca	média-pouca
50	pouca	média	muita	média	média	média
51	pouca	média	muita	média	muita	média
52	pouca	média	muita	muita	pouca	média-pouca
53	pouca	média	muita	muita	média	média
54	pouca	média	muita	muita	muita	média-muita

Regras	M Nível de Cinza	Tamanho	Orientação	Forma	Distância	Saliência Bottom-up
55	pouca	muita	pouca	pouca	pouca	média-pouca
56	pouca	muita	pouca	pouca	média	média-pouca
57	pouca	muita	pouca	pouca	muita	média
58	pouca	muita	pouca	média	pouca	média-pouca
59	pouca	muita	pouca	média	média	média
60	pouca	muita	pouca	média	muita	média
61	pouca	muita	pouca	muita	pouca	média-pouca
62	pouca	muita	pouca	muita	média	média
63	pouca	muita	pouca	muita	muita	média-muita
64	pouca	muita	média	pouca	pouca	média-pouca
65	pouca	muita	média	pouca	média	média-pouca
66	pouca	muita	média	pouca	muita	média
67	pouca	muita	média	média	pouca	média-pouca
68	pouca	muita	média	média	média	média
69	pouca	muita	média	média	muita	média
70	pouca	muita	média	muita	pouca	média
71	pouca	muita	média	muita	média	média
72	pouca	muita	média	muita	muita	média-muita
73	pouca	muita	muita	pouca	pouca	média-pouca
74	pouca	muita	muita	pouca	média	média
75	pouca	muita	muita	pouca	muita	média
76	pouca	muita	muita	média	pouca	média
77	pouca	muita	muita	média	média	média
78	pouca	muita	muita	média	muita	média-muita
79	pouca	muita	muita	muita	pouca	média
80	pouca	muita	muita	muita	média	média-muita
81	pouca	muita	muita	muita	muita	média-muita
82	média	pouca	pouca	pouca	pouca	pouca
83	média	pouca	pouca	pouca	média	média-pouca
84	média	pouca	pouca	pouca	muita	média-pouca
85	média	pouca	pouca	média	pouca	média-pouca
86	média	pouca	pouca	média	média	média-pouca
87	média	pouca	pouca	média	muita	média
88	média	pouca	pouca	muita	pouca	média-pouca
89	média	pouca	pouca	muita	média	média
90	média	pouca	pouca	muita	muita	média
91	média	pouca	média	pouca	pouca	média-pouca
92	média	pouca	média	pouca	média	média-pouca
93	média	pouca	média	pouca	muita	média
94	média	pouca	média	média	pouca	média-pouca
95	média	pouca	média	média	média	média
96	média	pouca	média	média	muita	média
97	média	pouca	média	muita	pouca	média-pouca
98	média	pouca	média	muita	média	média
99	média	pouca	média	muita	muita	média-muita
100	média	pouca	muita	pouca	pouca	média-pouca
101	média	pouca	muita	pouca	média	média-pouca
102	média	pouca	muita	pouca	muita	média
103	média	pouca	muita	média	pouca	média-pouca
104	média	pouca	muita	média	média	média
105	média	pouca	muita	média	muita	média
106	média	pouca	muita	muita	pouca	média
107	média	pouca	muita	muita	média	média
108	média	pouca	muita	muita	muita	média-muita
109	média	média	pouca	pouca	pouca	média-pouca
110	média	média	pouca	pouca	média	média-pouca

Regras	M Nível de Cinza	Tamanho	Orientação	Forma	Distância	Saliência Bottom-up
111	média	média	pouca	pouca	muita	média
112	média	média	pouca	média	pouca	média-pouca
113	média	média	pouca	média	média	média
114	média	média	pouca	média	muita	média
115	média	média	pouca	muita	pouca	média-pouca
116	média	média	pouca	muita	média	média
117	média	média	pouca	muita	muita	média-muita
118	média	média	média	pouca	pouca	média-pouca
119	média	média	média	pouca	média	média
120	média	média	média	pouca	muita	média
121	média	média	média	média	pouca	média-pouca
122	média	média	média	média	média	média
123	média	média	média	média	muita	média-muita
124	média	média	média	muita	pouca	média
124	média	média	média	muita	média	média
126	média	média	média	muita	muita	média-muita
127	média	média	muita	pouca	pouca	média-pouca
128	média	média	muita	pouca	média	média
129	média	média	muita	pouca	muita	média-muita
130	média	média	muita	média	pouca	média
131	média	média	muita	média	média	média
132	média	média	muita	média	muita	média-muita
133	média	média	muita	muita	pouca	média
134	média	média	muita	muita	média	média-muita
135	média	média	muita	muita	muita	muita
136	média	muita	pouca	pouca	pouca	média-pouca
137	média	muita	pouca	pouca	média	média
138	média	muita	pouca	pouca	muita	média
139	média	muita	pouca	média	pouca	média
140	média	muita	pouca	média	média	média
141	média	muita	pouca	média	muita	média-muita
142	média	muita	pouca	muita	pouca	média
143	média	muita	pouca	muita	média	média-muita
144	média	muita	pouca	muita	muita	média-muita
145	média	muita	média	pouca	pouca	média
146	média	muita	média	pouca	média	média
147	média	muita	média	pouca	muita	média-muita
148	média	muita	média	média	pouca	média
149	média	muita	média	média	média	média-muita
150	média	muita	média	média	muita	média-muita
151	média	muita	média	muita	pouca	média
152	média	muita	média	muita	média	média-muita
153	média	muita	média	muita	muita	muita
154	média	muita	muita	pouca	pouca	média
155	média	muita	muita	pouca	média	média
156	média	muita	muita	pouca	muita	média-muita
157	média	muita	muita	média	pouca	média
158	média	muita	muita	média	média	média-muita
159	média	muita	muita	média	muita	muita
160	média	muita	muita	muita	pouca	média-muita
161	média	muita	muita	muita	média	média-muita
162	média	muita	muita	muita	muita	muita
163	muita	pouca	pouca	pouca	pouca	média-pouca
164	muita	pouca	pouca	pouca	média	média-pouca
165	muita	pouca	pouca	pouca	muita	média
166	muita	pouca	pouca	média	pouca	média-pouca

Regras	M Nível de Cinza	Tamanho	Orientação	Forma	Distância	Saliência Bottom-up
167	muita	pouca	pouca	média	média	média
168	muita	pouca	pouca	média	muita	média
169	muita	pouca	pouca	muita	pouca	média
170	muita	pouca	pouca	muita	média	média
171	muita	pouca	pouca	muita	muita	média-muita
172	muita	pouca	média	pouca	pouca	média-pouca
173	muita	pouca	média	pouca	média	média
174	muita	pouca	média	pouca	muita	média
175	muita	pouca	média	média	pouca	média
176	muita	pouca	média	média	média	média
177	muita	pouca	média	média	muita	média-muita
178	muita	pouca	média	muita	pouca	média
179	muita	pouca	média	muita	média	média-muita
180	muita	pouca	média	muita	muita	média-muita
181	muita	pouca	muita	pouca	pouca	média-pouca
182	muita	pouca	muita	pouca	média	média
183	muita	pouca	muita	pouca	muita	média-muita
184	muita	pouca	muita	média	pouca	média
185	muita	pouca	muita	média	média	média-muita
186	muita	pouca	muita	média	muita	média-muita
187	muita	pouca	muita	muita	pouca	média
188	muita	pouca	muita	muita	média	média-muita
189	muita	pouca	muita	muita	muita	muita
190	muita	média	pouca	pouca	pouca	média-pouca
191	muita	média	pouca	pouca	média	média
192	muita	média	pouca	pouca	muita	média-muita
193	muita	média	pouca	média	pouca	média
194	muita	média	pouca	média	média	média
195	muita	média	pouca	média	muita	média-muita
196	muita	média	pouca	muita	pouca	média
197	muita	média	pouca	muita	média	média-muita
198	muita	média	pouca	muita	muita	muita
199	muita	média	média	pouca	pouca	média
200	muita	média	média	pouca	média	média
201	muita	média	média	pouca	muita	média-muita
202	muita	média	média	média	pouca	média
203	muita	média	média	média	média	média-muita
204	muita	média	média	média	muita	média-muita
205	muita	média	média	muita	pouca	média
206	muita	média	média	muita	média	média-muita
207	muita	média	média	muita	muita	muita
208	muita	média	muita	pouca	pouca	média
209	muita	média	muita	pouca	média	média-muita
210	muita	média	muita	pouca	muita	média-muita
211	muita	média	muita	média	pouca	média
212	muita	média	muita	média	média	média-muita
213	muita	média	muita	média	muita	muita
214	muita	média	muita	muita	pouca	média-muita
215	muita	média	muita	muita	média	muita
216	muita	média	muita	muita	muita	muita
217	muita	muita	pouca	pouca	pouca	média
218	muita	muita	pouca	pouca	média	média
219	muita	muita	pouca	pouca	muita	média-muita
220	muita	muita	pouca	média	pouca	média
221	muita	muita	pouca	média	média	média-muita
222	muita	muita	pouca	média	muita	muita

Regras	M Nível de Cinza	Tamanho	Orientação	Forma	Distância	Saliência Bottom-up
223	muita	muita	pouca	muita	pouca	média-muita
224	muita	muita	pouca	muita	média	média-muita
225	muita	muita	pouca	muita	muita	muita
226	muita	muita	média	pouca	pouca	média
227	muita	muita	média	pouca	média	média-muita
228	muita	muita	média	pouca	muita	muita
229	muita	muita	média	média	pouca	média-muita
230	muita	muita	média	média	média	média-muita
231	muita	muita	média	média	muita	muita
232	muita	muita	média	muita	pouca	média-muita
233	muita	muita	média	muita	média	muita
234	muita	muita	média	muita	muita	muita
235	muita	muita	muita	pouca	pouca	média-muita
236	muita	muita	muita	pouca	média	média-muita
237	muita	muita	muita	pouca	muita	muita
238	muita	muita	muita	média	pouca	média-muita
239	muita	muita	muita	média	média	muita
240	muita	muita	muita	média	muita	muita
241	muita	muita	muita	muita	pouca	média-muita
242	muita	muita	muita	muita	média	muita
243	muita	muita	muita	muita	muita	muita

A figura 5.56 simula objetos presentes em uma esteira. Os índices obtidos para os objetos (na seqüência da esquerda para a direita), foram respectivamente: 0.448, **0.769**, 0.401, 0.427, 0.403, **0.623**, 0.433. A chave de cano chamou mais a atenção porque apresenta uma forma diferente de todas as outras presentes na cena, entretanto todas as outras características de conspicuidade individuais interferem nos resultados obtidos para todos os objetos, resultando em um efeito global da cena no índice de cada objeto. A chave inglesa inclinada também chama a atenção, podendo indicar por exemplo o mal posicionamento de uma peça em uma esteira

Figura 5.56 – Exemplo de cena com objetos que apresentam diferentes características. O índice de saliência *bottom-up* determinado para cada objeto leva em consideração a influência de todas as características individuais em relação à todas as características dos outros objetos presentes na cena, resultando um efeito global da cena.

A figura 5.57 mostra um outro exemplo de cena onde os objetos compartilham diferentes características. Os índices obtidos para os objetos (na sequência indicada na cena) foram respectivamente: 0.634, 0.525, 0.653, **0.903**, 0.569, 0.569, 0.497.

Figura 5.57 – Exemplo de cena onde os objetos compartilham diferentes características. O objeto central foi o que obteve o índice de saliência *bottom-up* mais alto.

Na figura 5.58 a seguir, os índices obtidos para os objetos (na sequência indica na figura) foram respectivamente: **0.918**, 0.489, 0.432, 0.574, 0.455, **0.751**, 0.445.

Figura 5.58 – Exemplo de cena contendo objetos com diferentes características.

A figura 5.59 ilustra um exemplo onde todos os objetos apresentam pouca conspicuidade em relação aos outros (primeira regra). Na obtenção dos índices não foi

considerado um valor de *threshold* dentro das matrizes de saliências. Os resultados obtidos para os objetos foram: 0.174, 0.116, 0.0816, 0.0818, 0.182, 0.0965, 0.0848.

Figura 5.59 – Cena contendo objetos com pouca conspicuidade entre si.

A figura 5.60 ilustra uma situação onde um objeto difere muito dos demais objetos em todas as dimensões de características. O índice de saliência *bottom-up* obtido para o alicate foi 0.92 enquanto que os demais objetos apresentaram índices inferiores a 0.2.

Figura 5.60 – Exemplo de uma cena contendo um objeto que difere dos outros objetos em todas as dimensões de características.

A figura 5.61 a seguir mostra uma cena onde dois objetos se destacam em relação aos demais principalmente em relação a orientação e média de nível de cinza. Como o sistema foi projetado para apresentar uma maior relevância para a diferença de tonalidade em relação a inclinação, o objeto mais escuro apresentou um índice de 0.75, enquanto que o objeto com inclinação diferente apresentou um índice de 0.572. Os demais objetos apresentaram índices em torno de 0.25.

Figura 5.61 – Cena contendo um objeto com maior conspicuidade de média de nível de cinza e outro objeto com conspicuidade de orientação em relação aos demais objetos presentes.

A figura 5.62 mostra uma cena contendo alfinetes e pregos. Os índices obtidos para os pregos ficaram em torno de 0.5, enquanto que para os alfinetes (objetos 4 e 5) o índice de saliência *bottom-up* foi de 0.9.

Figura 5.62 – Cena contendo alfinetes e pregos.

2.7.2.3 – Inclusão de informações *Top-down*

Um procedimento *top-down* foi adicionado para guiar a atenção para um objeto específico na cena. Por definição a atenção *top-down* está relacionada com conhecimento de alto-nível, dependente da tarefa, em termos de quais objetos devem ser localizados ou quais características devem ser consideradas para localizá-los. Essas características podem ser baseadas nas propriedades dos objetos tais como orientação, cor, brilho, alongação, simetria, tamanho, bem como relações entre os objetos tais como menor (maior) que, à esquerda (à direita) de, abaixo (acima) de, mais claro (escuro) que, mais longe (perto), ou qualquer outra especificada pelo conhecimento humano.

Esses componentes entretanto dependem de uma aplicação específica. Nesse contexto, nos sistemas de visão computacional, assume-se que esse tipo de conhecimento está armazenado de alguma maneira no sistema de reconhecimento do objeto. Pode-se assumir que o sistema é capaz de analisar a imagem de entrada (*foco de atenção*) e usar seus modelos e conhecimentos pré-estabelecidos para contribuir na definição de localizações de interesse. Isso pode ser feito através da construção de um mapa de características *top-down*, cujos valores $T_{x,y}$ representam um grau de interesse da localização (x,y) , de acordo com o conhecimento sobre os objetos alvos ou conhecimento sobre as relações espaciais entre os objetos. Também podem ser definidos os valores apropriados para os parâmetros dos processos de integração dos mapas de características. Por exemplo, um determinado canal pode ter um peso mais alto se for conhecido ser uma característica importante para a extração de um objeto em particular.

O subsistema *fuzzy top-down* combina diferentes características de acordo com a relevância das mesmas na realização de diferentes tarefas, retornando um índice de saliência *top-down* em sua saída para cada objeto presente na cena. O resultado parcial para cada tarefa específica será combinado com o índice de saliência *bottom-up* no subsistema *fuzzy* de integração.

Novamente, a lógica *fuzzy* permite ao sistema tomar decisões com grande flexibilidade, através das regras de inferência. Utilizando esta estratégia, é possível especificar um conjunto de regras pré-definidas, para cada tarefa específica. As tarefas estão armazenadas em uma base de conhecimento. Devido ao uso de cinco características diferentes, o sistema é capaz de realizar oitenta e duas tarefas diferentes.

O subsistema *top-down*, orientado à tarefa, foi construído com 252 regras para cada tarefa através da combinação dos cinco conjuntos *fuzzy* de saída com os conjuntos *fuzzy* de entrada desse subsistema. Os conjuntos de entrada “Gray” (média de nível de cinza) e “tamanho” apresentam os termos lingüísticos (“pouco”, “médio” e “muito”) para as suas funções de pertinência e os conjuntos “orientação”, “forma” e “distância” apresentam os termos lingüísticos “pouco” e “muito”.

2.7.2.4 – Determinação das funções de pertinência do subsistema *top-down*

Para se obter as funções de pertinência para os conjuntos de entrada desse subsistema também foram realizadas pesquisas a fim de se extrair a percepção e conhecimento *top-down* humano.

Para a realização da pesquisa, foram entrevistadas cinquenta pessoas dentre elas especialistas da área de visão e processamento de imagens incluindo outros profissionais com formação superior em psicologia, comunicação visual, propaganda e marketing, fotografia, etc.

O objetivo desses testes psicofísicos, feito a partir de um questionário, foi extrair o conhecimento humano em questões subjetivas que expressem qualidades de:

- 1) Média de nível de cinza (Claro, tonalidade média, escuro)
- 2) Tamanho (pequeno, médio e grande)
- 3) Orientação (pouco, médio e muito inclinado)
- 4) Distância (perto, distância média, longe)

Os resultados obtidos com esse teste serviram para a formação da base de funções dos conjuntos *fuzzy* “nível de cinza”, “tamanho”, “orientação” e “distância” respectivamente, no subsistema *top-down*. As funções de pertinência determinam o grau de pertinência de uma determinada variável dentro do conjunto *fuzzy* correspondente e são usadas para aproximar a percepção humana, entretanto outros tipos de pesquisas podem ser utilizadas no modelamento das funções, de acordo com a aplicação e objetivo do sistema. Esse tipo de teste é *top-down* pois foram dadas pistas para as pessoas quanto as qualidades a serem quantificadas, como claro, escuro, perto, longe, inclinado, grande, pequeno, etc.

O conjunto de saída *fuzzy* “saliência *topdown*” é representado pelos seguintes termos lingüísticos: “pouca” (p), “média-pouca” (mp), “média”(m), “média-muita” (mg) e “muita” (g). A base de regras foi criada para armazenar diferentes regras tais como “pegar o menor alicate”, ou “pegar a chave inglesa mais escura” e assim por diante.

2.7.2.4.1 - Determinação das funções de pertinência para o conjunto *Fuzzy* “nível de cinza” do subsistema *top-down*

Na primeira parte do questionário, para se determinar as funções de pertinência do conjunto *fuzzy* “nível de cinza” foi usada a referência de níveis de cinza mostrada na figura 5.63. As pessoas investigadas foram questionadas a responder *quanto* da faixa de cinza atrai a atenção para as características subjetivas “tonalidade escura”, “tonalidade média” e “tonalidade clara”. Essas características subjetivas irão determinar as funções de pertinência para os termos “pouco”, “médio” e “muito” do conjunto *fuzzy Gray*.

Figura 5.63 – Faixa de nível de cinza compreendendo os valores de 0 a 255, utilizada como referência para a determinação das funções de pertinências para os termos “pouco”, “médio” e “muito” do conjunto *fuzzy* “nível de cinza”.

A fim de normalizar as respostas obtidas para cada função de pertinência, a faixa total de nível de cinza foi dividida em 10 subníveis, sendo o nível 1 definido para a faixa de tonalidade mais escura e o nível 10 definido para a tonalidade mais clara. As respostas obtidas a partir do questionário, com os respectivos níveis, são mostradas nas tabelas 5.24, 5.25 e 5.26.

Tabela 5.24 – Respostas obtidas para a característica “escuro”

“Pouco nível de cinza” (escuro)	Porcentagem de votos
1	32 %
2	28 %
3	21%
4	9 %
5	7 %
6	2 %
7	1%
8	0%
9	0%
10	0 %

As avaliações dos dados obtidos no questionário para os termos “pouco nível de cinza” (“escuro”), “médio nível de cinza” (tonalidade média) e “muito nível de cinza” (“claro”) são mostradas nas figuras 5.64, 5.65 e 5.66 respectivamente.

Figura 5.64 – Gráfico resultante da avaliação do questionário para o termo “pouco nível de cinza” (“escuro”).

Tabela 5.25 – Respostas obtidas para a característica “tonalidade média”

Tonalidade média	Porcentagem de votos
3	2 %
4	8 %
5	28%
6	28%
7	22%
8	12 %

Figura 5.65 – Gráfico resultante da avaliação do questionário para o termo “Médio Nível de Cinza”(tonalidade média).

Tabela 5.26 – Respostas obtidas para a característica “claro”

Muito nível de cinza (claro)	Porcentagem de votos
6	7,5%
7	13 %
8	18,5%
9	30%
10	31%

Figura 5.66 – Gráfico resultante da avaliação do questionário para o termo “*Large Gray*” (claro)

A figura 5.67 mostra a representação gráfica da três funções de pertinência do conjunto *fuzzy* “nível de cinza” construídas a partir das respostas processadas e a figura 5.68 mostra as funções que melhor representam as curvas obtidas.

Figura 5.67 – Representação gráfica das funções de pertinência (*membership functions*) do conjunto *fuzzy* “nível de cinza” (“gray”) para os termos “pouco” (“small”), “médio” (“medium”) e muito (“large”).

Figura 5.68 – Funções de pertinência para o conjunto *fuzzy* “nível de cinza” (“Gray”).

2.7.2.4.2 – Determinação das funções de pertinência para o conjunto *Fuzzy* “Tamanho” do subsistema *top-down*

Para se determinar as funções de pertinência do conjunto *fuzzy* “tamanho”, utilizou-se como referência de tamanho figuras análogas à mostrada na figura 5.69. As pessoas investigadas foram questionadas a responder *quanto* atrai a atenção para as características de tamanho: pequeno (atributo *pouco tamanho*), tamanho médio (atributo *tamanho médio*) e grande (atributo *muito tamanho*).

Os objetos dentro da figura foram normalizados em 10 tamanhos (tamanhos de 1 a 10), com o objeto menor e maior possuindo tamanhos “1” e “10” respectivamente, a fim de se determinar as funções de pertinência do conjunto *fuzzy* “Size”. As respostas obtidas a partir do questionário são mostradas nas tabelas 5.27, 5.28 e 5.29.

Figura 5.69 - Figura compreendendo objetos com tamanhos normalizados, utilizada como referência para a determinação das funções de pertinências para os atributos “pequeno”, “médio” e “grande” do conjunto *fuzzy* “tamanho”.

Tabela 5.27 – Respostas obtidas para a característica “Pequeno”

“Pouco Tamanho” (pequeno)	Porcentagem de votos
1	33 %
2	30 %
3	22%
4	11%
5	4%
6	0%
7	0%
8	0%
9	0%
10	0%

A avaliação dos dados obtidos na pesquisa para o termo “pequeno” é mostrado na figura 5.70.

Figura 5.70 - Gráfico resultante para o termo “pequeno” do conjunto *fuzzy* tamanho.

Tabela 5.28 – Respostas obtidas para a característica “tamanho médio”

“Tamanho médio”	Porcentagem de votos
3	12 %
4	23 %
5	31 %
6	23 %
7	9 %
8	2 %

A avaliação dos dados obtidos no questionário para o termo “tamanho médio” está mostrado na figura 5.71.

Figura 5.71 - Gráfico resultante da avaliação do questionário para o termo “tamanho médio”.

Tabela 5.29 – Respostas obtidas para a característica “Tamanho Grande”

“Tamanho grande”	Porcentagem de votos
6	2 %
7	8 %
8	20 %
9	33 %
10	37 %

A avaliação dos dados obtidos no questionário para o termo “tamanho grande” pode ser visto na figura 5.72.

Figura 5.72 - Gráfico resultante da avaliação do questionário para o termo “tamanho grande”.

A figura 5.73 mostra a representação gráfica da três funções de pertinência do conjunto *fuzzy* “Tamanho” construídas a partir das respostas processadas.

Figura 5.73 – Representação gráfica das funções de pertinência do conjunto *fuzzy* “Tamanho” para os termos “pequeno” (*small*), “médio” (*medium*) e muito (“*large*”).

A figura 5.74 mostra as funções que melhor representam as curvas obtidas com os resultados experimentais.

Figura 5.74 – Funções de pertinência para o conjunto *fuzzy* “Tamanho”.

2.7.2.4.3 – Determinação das funções de pertinência para o conjunto *Fuzzy* “Orientação” do subsistema *top-down*

Para se determinar as funções de pertinência do conjunto *fuzzy* “orientação” utilizou-se como referência figuras análogas à mostrada na figura 5.75. As pessoas investigadas foram questionadas a responder *quanto* atrai a atenção para as características de orientação: pouco inclinado e muito inclinado, tomando-se como referência uma reta vertical (0°). Os objetos da figura abaixo estão normalizados em seis orientações diferentes, onde a orientação 1 corresponde à 0° (reta vertical), e a orientação 6 corresponde à 90° (reta horizontal). Nessa experiência não foi considerado o sentido das retas (horário ou anti-horário) em relação a 0° .

Figura 5.75 - Figura compreendendo objetos com tamanhos normalizados, utilizada como referência para a determinação das funções de pertinências para os atributos “pouco” e “muito” do conjunto *fuzzy* “orientação”.

As respostas obtidas a partir do questionário são mostradas nas tabelas 5.30 e 5.31.

Tabela 5.30 – Respostas obtidas para a característica “pouco inclinado”

“Pouco Inclinado”	Porcentagem de votos
1	40 %
2	45 %
3	15%
4	0%
5	0%
6	0%

A figura 5.76 mostra a avaliação dos dados obtidos no questionário para o atributo “pouco inclinado” .

Figura 5.76 - Gráfico resultante da avaliação do questionário para o termo “pouco inclinado”

Tabela 5.31 – Respostas obtidas para a característica “muito inclinado”

“Muito Inclinado”	Porcentagem de votos
3	10 %
4	22 %
5	33 %
6	35 %

A avaliação dos dados obtidos no questionário para o termo “muito inclinado” é mostrada na figura 5.77.

Figura 5.77 - Gráfico resultante da avaliação do questionário para o termo “muito inclinado”.

A figura 5.78 mostra a representação gráfica das funções de pertinência do conjunto fuzzy “Orientação” construídas a partir das respostas processadas.

Figura 5.78 - Representação gráfica das funções de pertinência (*membership functions*) do conjunto fuzzy “Orientação” para os atributos “pouco” (*small*) e muito (“*large*”) inclinado.

A figura 5.79 mostra as funções que melhor representam as curvas obtidas com os resultados experimentais.

Figura 5.79 – Funções de pertinência aproximadas para o conjunto *fuzzy* “Orientação”

2.7.2.4.4 – Determinação das funções de pertinência para o conjunto *Fuzzy* “Distância” do subsistema *top-down*

Para se determinar as funções de pertinência do conjunto *fuzzy* “Distância” utilizou-se como referência de distância, figuras análogas à mostrada na figura 5.80. As pessoas investigadas foram questionadas a responder *quanto* atrai a atenção para as características de tamanho: perto (*pouca distância*), média distância e longe (*muita distância*), considerando o centro da figura como referência.

Os distâncias entre os objetos em relação ao centro da figura também foram normalizados em 10 valores, a fim de se determinar as funções de pertinência do conjunto *fuzzy* “Distância”. As respostas obtidas a partir do questionário são mostradas nas tabelas 5.32, 5.33 e 5.34.

A avaliação dos dados obtidos no questionário para a determinação da função de pertinência para “média distância” é mostrada na figura 5.82.

Figura 5.80 - Figura compreendendo objetos com tamanhos normalizados, utilizada como referência para a determinação das funções de pertinências para os atributos “pouca”, “média” e “muita” do conjunto *fuzzy* “Distância”.

Tabela 5.32 – Respostas obtidas para a característica “Perto”

“Pouca Distância”	Porcentagem de votos
1	28 %
2	25 %
3	23 %
4	9 %
5	9 %
6	6%
7	0%
8	0%
9	0%
10	0%

A avaliação dos dados obtidos no questionário para a determinação da função de pertinência para o atributo “perto” é mostrada na figura 5.81.

Figura 5.81 - Gráfico resultante da avaliação do questionário para o termo “Pouca Distância” (perto).

Tabela 5.33 – Respostas obtidas para a característica “média distância”

“Média Distância”	Porcentagem de votos
2	7 %
3	7 %
4	18 %
5	26 %
6	23 %
7	16 %
8	3 %

A avaliação dos dados obtidos no questionário para a determinação da função de pertinência para “média distância” é mostrada na figura 5.82.

Figura 5.82 - Gráfico resultante da avaliação do questionário para “Média Distância”

Tabela 5.34 – Respostas obtidas para a característica “longe”

“Muita Distância” (“Longe”)	Porcentagem de votos
5	1,5 %
6	1,5 %
7	18 %
8	23 %
9	28 %
10	28 %

A Avaliação dos dados obtidos no questionário para a função de pertinência para “Longe” (“muita distância”) está mostrada na figura 5.83

Figura 5.83 – Gráfico resultante da avaliação do questionário para “Longe” (“muita distância”).

A figura 5.84 mostra a representação gráfica das três funções de pertinência do conjunto *fuzzy* “Distância” construídas a partir das respostas processadas.

Figura 5.84 - Representação gráfica das funções de pertinência (*membership functions*) para os termos “pouca” (*small*), “média” (*medium*) e muita (*large*) distância.

A figura 5.85 mostra as funções que melhor representam as curvas obtidas com os resultados experimentais para caracterizar as funções de pertinência para os atributos “perto” (*small distance*), “distância média” (*medium distance*) e “longe” (*large distance*).

Figura 5.85 – Funções de pertinência para o conjunto *fuzzy* “Distância”

2.7.2.4.5 – Determinação das funções de pertinência para o conjunto *fuzzy* “Forma” do subsistema *top-down*

No subsistema *fuzzy top-down*, as funções de pertinência (*membership functions*) para o conjunto *fuzzy* “forma” representam os graus de pertinência para as ferramentas (cir, cci, ali, ccc, chb, chi, hll), conforme mostra a figura 5.86. Essas funções de pertinências foram derivadas das curvas de distribuição Gaussiana para a soma dos invariantes de momentos e estão mostradas na figura 5.10 e representa a principal característica para representar a forma dos objetos dentro da cena.

Figura 5.86 - Inferência chave para as tarefas envolvendo formas no sub-sistema *top-down*.

2.7.2.5 – Determinação das Regras de Inferência *fuzzy* para o subsistema *top-down*

O subsistema *top-down* combina as características de média de nível de cinza, tamanho, orientação, forma e distância, através das regras de inferência pré-definidas para cada tarefa específica. A relevância de cada característica na realização de uma determinada tarefa depende do comportamento do sistema que se deseja modelar e da precisão desejada. O subsistema *top-down*, orientado à tarefa, foi elaborado com 252 regras para cada tarefa através da combinação das cinco características de entrada e suas respectivas funções de pertinência, resultando em um índice de saliência *top-down* para cada objeto de acordo com a importância dos mesmos na realização da tarefa específica. Na elaboração das regras considerou-se apenas os termos lingüísticos (“pouca”, “média”, “muita”) para os conjuntos *fuzzy* de entrada “Cinza”, “Tamanho” e os termos (“pouca” e “muita”) para os conjuntos *fuzzy* “orientação” e “distância”. O conjunto *fuzzy* de entrada “forma” é composto pelas funções de pertinência dos objetos definidos para este sistema, de acordo com a distribuição dos valores da característica formato para cada objeto.

Na figura 5.87 podem ser vistas as funções de pertinência do conjunto *fuzzy* de saída “saliência *top-down*” que determinam o índice de saliência *topdown* após terem sido combinadas as informações dos cinco conjuntos *fuzzy* de entrada.

Figura 5.87 – Funções de pertinência para as variáveis “pouca” (p), “média-pouca” (mp), “média” (m), “média-muita” (mg) e “muita” (g) do conjunto *fuzzy* de saída “saliência *top-down*”.

O índice de saliência *top-down* (IST) é obtido através de um processo de “desfuzificação” que agrega as saídas obtidas para os conjuntos *fuzzy* de entrada, através do método da centróide.

A tabela A4.1 do apêndice 4 mostra como foram construídas as 252 regras de inferência utilizadas para determinar o índice de saliência *top-down* (IST) para a tarefa de “determinar os objetos escuros”.

A figura 5.88 mostra um exemplo de uma cena onde foram calculados os índices de saliência *top-down* para os objetos para a tarefa de determinar os objetos escuros. Os índices obtidos para os objetos 1 a 7 foram respectivamente: 0.583, 0.475, **0.91**, 0.556, 0.331, **0.907**, **0.737**.

Figura 5.88 – Cena contendo objetos com várias tonalidades.

A tabela A4.2 do apêndice 4 mostra como foram construídas as 252 regras de inferência utilizadas para determinar o índice de saliência *top-down* (IST) para a tarefa de “determinar os objetos grandes”.

A figura 5.89 mostra um exemplo de uma cena onde foram calculados os índices de saliência *top-down* para os objetos para a tarefa de “determinar os objetos grandes”. Os índices obtidos para os objetos 1 a 7 foram respectivamente: 0.0979, **0.907**, 0.105, 0.0918, 0.0985, 0.5, 0.51.

Figura 5.89 – Cena contendo objetos de diferentes tamanhos.

A figura 5.90 mostra um exemplo de uma cena onde foram calculados os índices de saliência *top-down* para os objetos para a tarefa de “determinar os círculos”. Os índices obtidos para os objetos 1 a 7 foram respectivamente: 0.387, 0.113, 0.113, 0.0906, 0.1, **0.892**, 0.0931.

A tabela A4.3 do apêndice 4 mostra como foram construídas as 252 regras de inferência utilizadas para determinar o índice de saliência *top-down* (IST) para a tarefa de “determinar os círculos”.

Na figura 5.89 os índices de saliência *top-down* para os objetos para a tarefa de “determinar os alicates”. Os índices obtidos para os objetos 1 a 7 foram respectivamente: **0.756**, 0.0954, 0.0918, **0.628**, **0.745**, 0.112, **0.792**.

Figura 5.90 – Cena contendo objetos com diferentes formatos.

As tarefas mostradas anteriormente são exemplos de tarefas que envolvem apenas uma dimensão de característica (média de nível de cinza, tamanho, orientação, forma ou distância). A tabela A4.4 do apêndice 4 mostra como foram construídas as 252 regras de inferência utilizadas para determinar o índice de saliência *top-down* (IST) para a tarefa de “determinar os círculos grandes” que combinam as características de forma e tamanho. A importância de uma determinada característica sobre as outras depende de como se deseja modelar a resposta do sistema. Quando a forma de um objeto está envolvida com outras características, a característica de formato neste sistema apresenta uma maior relevância. Conforme pode ser visto na tabela A4.4, no caso da tarefa de determinar os círculos grandes, os círculos de tamanho grande, médio e pequeno apresentaram uma maior importância nesta ordem, seguida por quaisquer outro objeto de tamanho grande e por último estão os outros objetos de outros tamanhos que despertam pouca atenção na execução da tarefa.

A figura 5.91 ilustra um exemplo de uma cena onde foram calculados os índices de saliência *top-down* para os objetos para a tarefa de “determinar os círculos grandes”. Os índices obtidos para os objetos 1 a 7 foram respectivamente: 0.0833, **0.756**, **0.625**, 0.25, **0.5**, 0.091, **0.5**. A chave inglesa (Chb) apresentou um índice de saliência *top-down* maior que os demais objetos devido ao seu tamanho.

Figura 5.91 – Cena contendo diferentes objetos com diferentes tamanhos e tonalidades.

A tabela A4.5 do apêndice 4 mostra como foram construídas as 252 regras de inferência utilizadas para determinar o índice de saliência *top-down* (IST) para a tarefa de “determinar os alicates escuros” que combinam as características de forma e média de nível de cinza.

Na figura 5.89 os índices de saliência *top-down* para os objetos para a tarefa de “determinar os alicates escuros”. Os índices obtidos para os objetos 1 a 7 foram respectivamente: **0.674**, 0.201, 0.193, **0.61**, **0.597**, 0.212, **0.693**. Conforme pode ser visto nos resultados, os alicates obtiverem um maior índice de saliência *top-down*. Os índices obtidos para os alicates dependeram da relação do nível de cinza com relação a tarefa. Os objetos diferentes do alicate apresentaram um *IST* maior em função de possuírem valores de média de nível de cinza mais próximos com relação a tarefa desejada.

As duas tarefas mostradas anteriormente exemplificaram uma situação onde a tarefa a ser realizada depende da análise de duas dimensões de características. A tabela A4.6 do apêndice 4 mostra como foram construídas as 252 regras de inferência utilizadas para determinar o índice de saliência *top-down* (IST) para a tarefa de “determinar os Grifos escuros e inclinados” que combinam as características de forma, média de nível de cinza e orientação. O critério utilizado prevaleceu a característica tonalidade escura em relação a orientação, conforme pode ser visto na tabela de regras A4.6.

A figura 5.92 ilustra um exemplo de uma cena onde foram calculados os índices de saliência *top-down* para os objetos para a tarefa de “determinar os grifos escuros e menos inclinados (com relação a vertical)”. Os índices obtidos para os objetos 1 a 7 foram respectivamente: **0.581**, **0.622**, 0.397, **0.631**, 0.0967, **0.56**, 0.0926. Os índices mais altos foram obtidos para os grifos. O objeto 3 apresentou um índice mais baixo principalmente devido a uma distorção em sua forma, fazendo com que a característica de formato computada apresentasse um valor bem inferior a média. O objeto 6 apesar de ser o grifo mais inclinado apresentou um índice inferior ao grifo escuro pouco inclinado.

Figura 5.92 – Cena contendo objetos (alicates e grifos) com diferentes tonalidades e inclinações.

A tabela A4.7 do apêndice 4 mostra como foram construídas as 252 regras de inferência utilizadas para determinar o índice de saliência *top-down* (IST) para a tarefa de “determinar a chave inglesa, clara, pequena e distante” que combina as características de forma, média de nível de cinza, tamanho e distância.

A figura 5.93 ilustra um exemplo de uma cena onde foi calculado os índices de saliência *top-down* para os objetos para a tarefa de “determinar a chave inglesa clara, pequena e mais distante”. Os índices obtidos para os objetos 1 a 7 foram respectivamente: 0.247, 0.0966, 0.459, 0.664, 0.248, 0.248, 0.685. Os índices mais altos foram obtidos para as chaves inglesas mais claras.

Figura 5.93 – Cena contendo objetos com diferentes características.

A tabela A4.8 do apêndice 4 mostra como foram construídas as 252 regras de inferência utilizadas para determinar o índice de saliência *top-down* (IST) para a tarefa de “determinar a chave hallen, escura, grande, mais inclinada e mais próxima” que combina as características de forma, média de nível de cinza, tamanho, orientação e distância.

A figura 5.94 ilustra um exemplo de uma cena onde foi calculado os índices de saliência *top-down* para os objetos para a tarefa “determinar a chave hallen, escura, grande, mais inclinada e mais próxima”. Os índices obtidos para os objetos 1 a 7 foram respectivamente: 0.5, 0.0109, 0.098, 0.648, 0.5, 0.25, 0.5.

Figura 5.94 – Cena contendo objetos com diferentes características.

2.7.2.6– O subsistema *Fuzzy* de Integração

O objetivo do subsistema de integração é analisar os índices de saliência obtidos nos subsistemas *bottom-up* e *top-down*, e fornecer um índice de saliência global. As regras usadas nesse subsistema foram especificadas a fim de aumentar ou inibir os resultados obtidos no sistema *bottom-up*, se uma tarefa específica concorda ou discorda com ele. O índice de saliência global permite também avaliar a influência do subsistema *bottom-up* no subsistema *top-down*.

A tabela 5.35 mostra como foram determinadas as 25 regras que relacionam os conjuntos *fuzzy* de entrada “índice de saliência *bottom-up*” e “índice de saliência *top-down*” com o conjunto *fuzzy* de saída “saliência global”. O índice de saliência *global* (*ISG*) é obtido através de um processo de “desfuzificação” que agrega as saídas obtidas para os conjuntos *fuzzy* de entrada, através do método da centróide.

Tabela 5.35 – Regras de Inferência para determinação do índice de Saliência Global”

REGRAS	ÍNDICE DE SALIÊNCIA TOP-DOWN	ÍNDICE DE SALIÊNCIA BOTTOM-UP	SALIÊNCIA GLOBAL
1	Pouca	Pouca	Pouca
2	Pouca	Média-Pouca	Pouca
3	Pouca	Média	Pouca
4	Pouca	Média-Muita	Pouca
5	Pouca	Muita	Média-Pouca
6	Média-Pouca	Pouca	Pouca
7	Média-Pouca	Média-Pouca	Média-Pouca
8	Média-Pouca	Média	Média-Pouca
9	Média-Pouca	Média-Muita	Média-Pouca
10	Média-Pouca	Muita	Média
11	Média	Pouca	Média-Pouca
12	Média	Média-Pouca	Média-Pouca
13	Média	Média	Média
14	Média	Média-Muita	Média
15	Média	Muita	Média-Muita
16	Média-Muita	Pouca	Média
17	Média-Muita	Média-Pouca	Média

REGRAS	ÍNDICE DE SALIÊNCIA TOP-DOWN	ÍNDICE DE SALIÊNCIA BOTTOM-UP	SALIÊNCIA GLOBAL
18	Média-Muita	Média	Média-muita
19	Média-Muita	Média-Muita	Média-muita
20	Média-Muita	Muita	Média-Muita
21	Muita	Pouca	Média-Muita
22	Muita	Média-Pouca	Média-Muita
23	Muita	Média	Média-Muita
24	Muita	Média-Muita	Muita
25	Muita	Muita	Muita

Na figura 5.95 podem ser vistas as funções de pertinência do conjunto *fuzzy* de saída “saliência *global*” que determinada o índice de *saliência global* após terem sido combinadas os índices provenientes dos dois subsistemas *fuzzy bottom-up* e *top-down*. As funções de pertinência para a variável de saída “saliência *global*” foram definidas para os atributos “pouca” (p), “média-pouca” (mp), média (m), “média-muita” (mg), e “muita”(g).

Figura 5.95 – Funções de pertinência para as variáveis “pouca” (p), “média-pouca” (mp), “média” (m), “média-muita” (mg) e “muita” (g) do conjunto *fuzzy* de saída “saliência *top-down*”.

A figura 5.96 mostra um exemplo de uma cena onde foram calculados os índices de saliência *bottom-up*, *top-down* e *global* para os objetos para na tarefa de “determinar o círculo grande” e na tarefa “determinar o círculo pequeno”. As tabelas 5.36 e 5.37 mostram os índices (*bottom-up*, *top-down* e *global*) obtidos para os objetos, para cada tarefa. Na figura 5.97 pode ser visto uma representação do mapa de saliência *bottom-up* medido para a cena mostrada na figura 5.96 e nas figuras 5.98 e 5.99 é mostrado uma comparação gráfica entre os índices medidos para a figura 5.96 para as duas tarefas realizadas.

Figura 5.96 – Cena com objetos utilizada na tarefa de “determinar o círculo grande” e na tarefa “determinar o círculo pequeno”.

Figura 5.97 – Mapas de Conspicuidades de Média de Nível de Cinza, tamanho, orientação, formato e distância calculados para os objetos da figura 5.96.

Tabela 5.36 – Comparação entre os índices de saliência *bottom-up*, *top-down* e *global*, obtidos para os objetos da figura 5.96 para a tarefa de determinar o “círculo pequeno”

OBJETO	ISB	IST	ISG
1	0,208	0,5	0,5
2	0,208	0,5	0,5
3	0,211	0,5	0,5
4	0,212	0,5	0,5
5	0,75	0,882	0,786
6	0,209	0,5	0,5
7	0,207	0,5	0,5

Tabela 5.37 – Comparação entre os índices de saliência *bottom-up*, *top-down* e *global* obtidos para os objetos da figura 5.97 para a tarefa de determinar o “círculo grande”.

OBJETO	ISB	IST	ISG
1	0,208	0,882	0,723
2	0,208	0,882	0,723
3	0,211	0,882	0,727
4	0,212	0,882	0,728
5	0,75	0,5	0,5
6	0,209	0,882	0,724
7	0,207	0,882	0,722

Figura 5.98 – Índices medidos para a cena mostrada na figura 5.96, para a tarefa de determinar os “círculos grandes”.

Figura 5.99 – Índices medidos para a cena mostrada na figura 5.96, para a tarefa de determinar os “círculos pequenos”.

Conforme mostrado nas tabelas 5.35 e 5.36, o índice de saliência *bottom-up* teve um efeito diferente sobre os objetos na realização das diferentes tarefas. O índice de saliência global para o objeto alvo na tarefa de pegar o círculo pequeno foi maior do que para a tarefa de determinar o círculo grande, devido a contribuição *bottom-up*. Na cena, todos os objetos apresentaram um índice mais alto pois para realizar a tarefa, foram consideradas nas regras uma maior importância para a forma do que para o tamanho, assim qualquer círculo apresenta uma saliência top-down média que cresce à medida que seu tamanho diminui (ou aumenta) em relação aos demais objetos presentes na cena (de acordo com a tarefa).

A figura 5.100 mostra uma cena análoga à mostrada na figura 5.96, mas o objeto alvo para a tarefa de determinar os círculos pequenos apresenta agora uma contribuição maior, pois sua forma é diferente de todos os objetos na cena.

Figura 5.100 – Cena análoga à mostrada na figura 5.96, onde o objeto alvo (círculo pequeno) apresenta uma contribuição *bottom-up* maior devido à diferença de formato com os demais objetos presentes na cena.

A figura 5.101 mostra os mapas de conspicuidade calculados para a cena mostrada na figura 5.100.

Figura 5.101 – Mapas de Conspicuidade calculados para os objetos presentes na figura 100.

A tabela 5.38 mostra os índices obtidos para os objetos da figura 5.100. A figura 5.102 mostra uma comparação gráfica dos índices medidos.

Tabela 5.38 – Comparação entre os índices de saliência *bottom-up*, *top-down* e *global*, obtidos para os objetos da figura 5.100 para a tarefa de determinar o “círculo pequeno”

OBJETO	ISB	IST	ISG
1	0,267	0,0875	0,194
2	0,267	0,113	0,211
3	0,275	0,0968	0,201
4	0,27	0,0875	0,194
5	0,92	0,756	0,812
6	0,288	0,0882	0,202
7	0,321	0,103	0,205

Figura 5.102 – Índices medidos para a cena mostrada na figura 5.100, para a tarefa de determinar os “círculos pequenos”.

O índice de saliência global obtido para o mesmo objeto alvo foi maior para a cena mostrada na figura 5.100 devido a maior influência de sua componente *bottom-up*.

A figura 5.103 mostra um outro exemplo de uma cena onde foram calculados os índices de saliência *bottom-up*, *top-down* e *global* para os objetos para a tarefa de determinar os círculos pequenos. A tabela 5.39 mostra os índices (*bottom-up*, *top-down* e *global*) obtidos para os objetos. Na figura 5.104 pode ser visto uma representação do mapa de saliência *bottom-up* medido para a cena mostrada na figura 5.103.

Figura 5.103 – Exemplo de uma cena para demonstrar como a presença de um objeto na cena pode interferir na execução de uma tarefa.

Figura 5.104 – Mapas de Conspicuidades de Média de Nível de Cinza, tamanho, orientação, formato e distância calculados para os objetos da figura 5.103.

Tabela 5.39 – Comparação entre os índices de saliência para os objetos da figura 5.103.

OBJETO	ISB	IST	ISG
1	0,918	0,112	0,234
2	0,248	0,756	0,74
3	0,288	0,882	0,786
4	0,28	0,882	0,786
5	0,279	0,9	0,8
6	0,245	0,769	0,744
7	0,283	0,882	0,786

Figura 5.105 – Comparação dos índices de saliência obtidos nos subsistemas *bottom-up*, *top-down* e de integração para a cena mostrada na figura 5.103.

Conforme pode ser visto na tabela 5.39 e na figura 5.105, o índice de saliência global para os círculos não foi reforçado pela componente *bottom-up* no processo de integração das características. Apenas a saliência do alicate apresentou um pequeno aumento devido a sua componente *bottom-up* ser muito alta. Todos os círculos apresentaram uma atividade significativa. A presença do alicate muito grande na cena em relação aos demais objetos na cena também interferiu na tarefa de detecção dos círculos pequenos, porque foi o tamanho utilizado como referência para normalização e por comparação, todos os demais objetos são “pequenos”.

Na figura 5.103, a tarefa de pegar o alicate pode ser favorecida pela componente *bottom-up* do alicate. A tabela 5.40 mostra os índices obtidos antes e após a integração.

Tabela 5.40 – Comparação entre os índices de saliência obtidos para os objetos da figura 5.103.

OBJETO	ISB	IST	ISG
1	0,918	0,829	0,82
2	0,248	0,121	0,216
3	0,288	0,121	0,216
4	0,28	0,121	0,216
5	0,279	0,121	0,216
6	0,245	0,121	0,216
7	0,283	0,121	0,216

A figura 5.106 mostra um outro exemplo envolvendo diversas tarefa e a figura 5.107 mostra os mapas de conspicuidade obtidos para os objetos da cena.

Figura 5.106 – Cena contendo chave inglesas e alicate.

Figura 5.107 – Mapa de conspicuidade obtido para os objetos da cena mostrada na figura 5.104.

A tabela 5.41 mostra os índices obtidos antes e após a integração para a tarefa de: (1)determinar o objeto escuro, (2)determinar o alicate, (3) determinar o alicate escuro , (4) determinar o grifo (chi) escuro e inclinado e (5) determinar a chave inglesa (chb) clara, menor e mais distante.

Tabela 5.41 – Índices obtidos para os objetos da figura 5.106, em diversas tarefas.

OBJ	ISB	IST1	IST2	IST3	IST4	IST5	ISG1	ISG2	ISG3	ISG4	ISG5
1	0,287	0,597	0,089	0,191	0,448	0,531	0,602	0,195	0,242	0,438	0,54
2	0,173	0,631	0,096	0,225	0,494	0,498	0,538	0,2	0,237	0,414	0,414
3	0,24	0,608	0,104	0,204	0,465	0,361	0,596	0,206	0,245	0,455	0,363
4	0,236	0,61	0,096	0,206	0,469	0,358	0,592	0,2	0,245	0,459	0,361
5	0,173	0,63	0,094	0,225	0,495	0,498	0,537	0,2	0,237	0,414	0,414
6	0,287	0,595	0,089	0,191	0,448	0,531	0,601	0,195	0,242	0,438	0,54
7	0,92	0,65	0,872	0,735	0,083	0,083	0,691	0,798	0,781	0,236	0,236

Conforme pode ser observado nos resultados, o objeto mais conspícuo (*bottom-up*) é o alicate. Na tarefa de determinar os objetos escuros, devido ao fato do alicate apresentar uma média de nível de cinza um pouco menor, o índice determinado para ele foi maior, sendo que a componente *bottom-up* serviu para realçar um pouco mais as diferenças entre os índices obtidos para os objetos na tarefa, após a integração.

A segunda tarefa que consiste em determinar os alicates escuros, o alicate apresentou um maior índice, através da realização da tarefa que envolve combinação de forma e cor. A tarefa de se determinar o alicate escuro e inclinado (terceira tarefa) apresentou para o alicate um índice menor do que o obtido na tarefa anterior, devido ao fato do objeto não possuir uma das características desejada (menos significativa), ou seja, não ser inclinado.

Na tarefa de pegar os grifos (quarta tarefa), devido a forma das funções de pertinência definida para o grifo e para a chave inglesa (ver figura 5.86), o sistema apresentou um índice mais elevado para a chave inglesa, que apesar de não ser o objeto alvo, assemelha-se a ele. Neste caso, pode ser estabelecido algum *threshold*, onde abaixo de um determinado valor, o objeto alvo em questão pode ser considerado não presente na cena. Na tarefa de determinar a chave inglesa mais clara, pequena e distante, os índices obtidos foram mais baixos pois os objetos presentes não compartilham todas as características com o alvo. Os maiores índices obtidos nessa tarefa foi para as chaves inglesas mais distante do centro de massa da cena.

Nas regras de pertinência para as tarefas que envolvem a combinação de forma com outras características, utilizou-se uma maior relevância para o formato dos objetos, assim por

exemplo quando deseja-se determinar os círculos grandes em uma cena, qualquer círculo na cena apresenta um potencial “médio” em ser o alvo, aumentando essa saliência à medida que o tamanho dos objetos aumenta. A variação do índice de saliência *top-down* depende das características das funções de pertinência dos conjunto *fuzzy* de entrada e saída. Se o valor de *threshold* adotado for 0.5 para a quarta e quinta tarefas, pode se presumir que o objeto alvo não está presente na cena no caso tarefa de determinar os grifos, e escolher apenas as chaves inglesas mais distantes no caso da tarefa de determinar a chave inglesa, clara, pequena e inclinada.

A figura 5.108 é análoga à cena mostrada na figura 5.106 com exceção do grifo. O mapa de conspicuidade calculado para a cena 5.108 pode ser visto na figura 5.109. Os índices obtidos para a tarefa de detectar o grifo escuro e inclinado pode ser visto na tabela 5.42.

Figura 5.108 – Cena análoga à mostrada na figura 5.106 onde o alvo (grifo) está presente.

Figura 5.109 – Mapas de Conspicuidade para a cena 5.108.

Tabela 5.42 – Índices obtidos para a cena mostrada na figura 5.108, para a tarefa de determinar o grifo, escuro e inclinado.

OBJETO	ISB	IST	ISG
1	0,254	0,448	0,438
2	0,281	0,494	0,49
3	0,333	0,465	0,45
4	0,254	0,469	0,46
5	0,391	0,495	0,388
6	0,905	0,608	0,671
7	0,754	0,083	0,19

A figura 5.110 ilustra um outro exemplo onde a tarefa consiste em se determinar os alicates escuros.

Figura 5.110 – Cena contendo alicates que compartilham diferentes características.

O mapa de conspicuidade calculado para a cena 5.110 pode ser visto na figura 5.111. Os índices obtidos para a tarefa de detectar o grifo escuro e inclinado pode ser visto na tabela 5.43.

Figura 5.111 – Mapa de conspicuidade para a cena mostrada na figura 5.110.

Tabela 5.43 – Índices obtidos para a cena mostrada na figura 5.110, para a tarefa de determinar o alicate escuro.

OBJETO	ISB	IST	ISG
1	0,854	0,735	0,749
2	0,582	0,502	0,503
3	0,582	0,502	0,502
4	0,584	0,548	0,56
5	0,878	0,735	0,758
6	0,91	0,588	0,671
7	0,742	0,527	0,536

De acordo com o processo *bottom-up*, os objetos mais conspícuos são os objetos 6, 5, 1 e 4 respectivamente nesta ordem. De acordo com a tarefa, os objetos mais salientes são o 1 e 2 , seguido pelo objeto 6. O objeto 6, apesar de ser o mais escuro, apresentou um índice *top-down* menor entre os alicates escuros, devido a sua diferença de formato (distorção) em relação aos demais pois a característica de forma é mais relevante que a de tonalidade na realização da tarefa.

A figura 5.112 a seguir mostra os índices de saliência obtidos para os objetos mostrados na figura 5.110.

Figura 5.112 – Índices de saliência *bottom-up*, *top-down* e *global*, obtidos para a tarefa de determinar os alicates escuros na cena mostrada na figura 5.110.

A figura 5.113 ilustra uma outra imagem contendo objetos com mesmo formato, compartilhando diferentes características. A tarefa nessa cena consiste em se determinar a chave inglesa clara, pequena e mais distante. Como o alvo não está presente na cena, os objetos apresentaram pouca saliência *topdown*, sendo que o objeto 6 foi o que mais se aproximou do objeto alvo, de acordo como foram determinadas as regras para realizar essa tarefa.

Figura 5.113 – Imagem contendo chaves inglesas com diferentes características.

Os mapas de conspicuidade para os objetos da cena mostrada na figura 5.113 podem ser vistos na figura 5.114. Na tabela 5.44 são mostrados os índices obtidos para a cena 5.113 para a tarefa de determinar a chave inglesa, clara, pequena e mais distante.

Figura 5.114 – Mapas de Conspicuidade medido para os objetos presentes na cena mostrada na figura 5.113.

Tabela 5.44 – Índices obtidos para a cena mostrada na figura 113, para a tarefa de determinar chave inglesa clara, pequena e mais distante

OBJETO	ISB	IST	ISG
1	0,435	0,25	0,25
2	0,414	0,25	0,25
3	0,698	0,273	0,301
4	0,479	0,25	0,25
5	0,89	0,25	0,25
6	0,76	0,25	0,25
7	0,625	0,25	0,25

2.8 – Discussão dos Resultados

Nas sessões anteriores a performance do modelo proposto foi exemplificada através de algumas imagens sintéticas e reais. Em várias cenas provenientes do subsistema *bottom-up*, as regiões de maior atenção corresponderam à algum tipo de irregularidade dos objetos dentro da cena. No subsistema *top-down*, as regiões mais interessantes corresponderam aos objetos alvos, quando os mesmos estavam presentes. Devido a metodologia proposta se basear em uma rede *fuzzy* de duas camadas, o índice de integração global (ISG) define melhor a performance do sistema como um todo, entretanto os índices de saliência *bottom-up* (ISB) e *top-down* (IST) fornecem informações interessantes sobre os objetos como por exemplo medidas de irregularidade entre os objetos e medida de semelhança dos objetos com o alvo.

A comparação dos índices ISB, IST e ISG dentro de uma mesma cena também fornecem dados importantes sobre a cena como um todo. A integração dos índices permite também avaliar a influência do subsistema *bottom-up* em realizar uma determinada tarefa. O índice de saliência ISB pode fortalecer, inibir ou manter o índice IST dependendo das características de interação entre os dois subsistemas definidas previamente. A saída do sistema de integração (ISG) entretanto, na maioria das situações, prevalece a determinação do alvo definido pelo subsistema *top-down* em relação ao *bottom-up*. Através da comparação de ISG com ISB e IST é possível analisar a intensidade da influência que o subsistema *bottom-up* exerce sobre o subsistema *top-down*. A comparação entre os índices para uma mesma cena ou em diferentes cenas considerando diferentes tarefas também indica um determinado grau de facilidade ou dificuldade que as componentes *bottom-up* e *top-down* da cena influenciam a atenção para os objetos dentro da cena.

Muitos parâmetros utilizados para modelar as funções de pertinência basearam-se em testes simples envolvendo a percepção humana. Outros parâmetros tais como as funções de pertinência para os conjuntos *fuzzy* “Saliência *Bottom-up*”, “Saliência *Top-down*” e “Saliência Global” e a ponderação das características dentro das regras dos três subsistemas *fuzzy* foram adotadas empiricamente, de acordo com a importância das mesmas em uma tarefa ou cena. Os ajustes destes parâmetros do sistema são fundamentais para a realização de uma função específica. No que concerne ao modelamento da percepção humana, a determinação de parâmetros mais precisos através do desenvolvimento de testes mais adequados e da aplicação dos mesmos em um maior número de pessoas, pode auxiliar no desenvolvimento de uma métrica para avaliar os diferentes graus de interferência entre as

características envolvidas nos processos *bottom-up* e *top-down*. Testes envolvendo combinações específicas de características onde o alvo não está presente, podem revelar a importância de uma determinada característica em relação a outras características e auxiliar na escolha correta dos pesos nas regras. Usando regras *fuzzy* é importante observar apenas quais regiões de interesse no domínio são adequadamente cobertas por espaços de conhecimentos apropriados, portanto os sistemas que usam regras nebulosas são mais simples de construir e manter (Politano, 1996).

Tem-se observado, em um grande número de aplicações comerciais, que as formas exatas dos conjuntos nebulosos não são muito críticas para o sucesso ou falha de um sistema nebuloso. Isso é devido a natureza especial do procedimento de inferência *fuzzy* (Dutta, 1993) apud (Politano, 1996). Os operadores “max-min” são os mais utilizados em aplicações comerciais devido a sua simplicidade.

O sistema foi testado com algumas imagens sugerindo possibilidades de aplicações industriais. Os resultados obtidos (índices de saliência) para as regiões de maior atenção corresponderam sempre ao objeto alvo (quando o mesmo estava presente) e foram capazes de tomar decisão e selecionar o melhor objeto em situações de conflito, onde existe interferências entre as características e o alvo não estava presente. Através da lógica *fuzzy* é possível modelar sistemas complexos que permitem a combinação diferenciada de diversas características. As respostas obtidas mostraram que o sistema é capaz de fornecer informações interessantes sobre os objetos e a sobre a cena global, realizar reconhecimento de padrões, combinar e ponderar diferentes características necessárias para realização de diferentes tarefas. O sistema permite através da lógica *fuzzy* a combinação de diferentes graus de interesse para os objetos, permitindo também meios de inferir quando o objeto alvo está ou não presente na cena.

CONCLUSÕES

1 – CONSIDERAÇÕES FINAIS

Uma componente importante das rotinas visuais comportamentais é a habilidade de determinar um item em um mundo visual composto de vários itens “distrativos”. A habilidade de realizar busca visual têm sido assunto de pesquisa principalmente nos últimos vinte anos. A facilidade com que os seres humanos encontram o alvo em um ambiente contendo diferentes objetos que compartilham diferentes características é também assunto desta tese. Essa habilidade não é trivial. O sistema visual não pode processar simultaneamente todas as entradas pois não possuímos espaço nem *hardware* neuronal para realizar todas as funções visuais em todas as localizações ao mesmo tempo (Tsotsos, 1988). Como sugerido por Neisser (1967), existem processos paralelos que operam sobre uma grande porção do campo visual de uma vez, e existe um segundo conjunto de processos de capacidade limitada que são restritos a pequenas porções do campo visual e são deslocados serialmente em localizações específicas. Esses processos não são controlados de maneira randômica e sim por mecanismos atencionais. Através da atenção é possível levar algum processo restrito a uma posição específica do campo visual. O controle dos mecanismos atencionais pode ser exógeno, baseado nas propriedades dos estímulos visuais, ou endógeno, baseado nas necessidades do “usuário” do sistema visual. Esses controles não são mutuamente exclusivos. O desenvolvimento de uma determinada ação pode ser baseada no desejo do observador em olhar para um objeto específico. Nesta tese nós consideramos o desenvolvimento da atenção endógena, exógena e mista.

Este trabalho prevê o desenvolvimento de uma metodologia de atenção visual, para operar em cenas simples, contendo um número limitado de objetos distribuídos aleatoriamente e abrange uma classe de tarefas envolvidas com objetos bidimensionais ou tridimensionais que possam ser projetados em duas dimensões mantendo entretanto suas características significativas.

A literatura apresentada nesta tese procurou integrar informações referentes aos diferentes aspectos da Atenção, Busca e Percepção visuais. A simulação apresentada aqui pode, com um simples ajuste de alguns parâmetros, ser adaptada para realizar uma grande variedade de tarefas visuais inclusive permite a inclusão de dados que se assemelhem à percepção humana.

A tarefa de busca visual de um objeto alvo entre um número de itens distrativos são mais fáceis para alguns estímulos do que para outros. Por exemplo, um alvo definido por uma única cor ou por uma única orientação é facilmente detectado (Treisman & Gelade, 1980). Muitos testes realizados com Busca Visual utilizam tempo de reação para avaliar a dificuldade de reconhecimento em tarefas onde o alvo está presente ou não. Wolfe, Cave & Franzel, (1989) mediram o tempo de reação envolvido na realização de diversas tarefas definidas por conjunções de cor e forma, conjunção de cor e orientação e conjunção de cor e tamanho.

A análise da complexidade computacional da percepção visual revela que os mecanismos de atenção são necessários a fim de realizar tarefas como busca visual e reconhecimento de objetos. O modelo proposto nesta tese para atenção visual foi obtido através de uma abordagem multi-disciplinar, que combinou sugestões obtidas de sistemas visuais biológicos e teorias de percepção humana, com uma análise computacional do problema. Em resumo, o modelo está estruturado nos seguintes estágios:

- Segmentação dos objetos em relação ao fundo através de uma técnica de Crescimento de Região.
- Representação da imagem através de múltiplos mapas de características independentes nos subsistemas *bottom-up* e *top-down* e de integração.
- A integração (combinação diferenciada) das características em todos os subsistemas é realizada através de uma abordagem *fuzzy* e levou em consideração características que atraem a atenção nos seres humanos.

A estrutura e a implementação proposta para os passos apresentaram uma série de vantagens, incluindo:

- Uso de esquemas computacionais distribuídos e paralelizados que permitem a implementação paralela.
- Possibilidade de estender o sistema através de detecções de regiões de interesse no tempo, permitindo também que o sistema possa controlar a interação entre os subsistemas de maneira dinâmica.

2 – EXTENSÃO DO MODELO BÁSICO E SUGESTÕES PARA TRABALHOS FUTUROS

O sistema atencional apresentado nesta tese é composto de três subsistemas. O subsistema *bottom-up* não utiliza nenhum tipo de conhecimento prévio sobre os objetos de interesse. A vantagem desse subsistema é que ele pode ser empregado em uma grande variedade de imagens e aplicações sem grandes mudanças na arquitetura, entretanto se algum tipo de conhecimento de alto nível estiver disponível, ele pode ser utilizado para aumentar a performance do sistema.

O subsistema *top-down* executa diversas tarefas e atua também como um sistema de reconhecimento de objeto, permitindo a combinação da forma com outras características. Este subsistema também produz um mapa de saliência *top-down* que pode ser integrado com os mapas de saliência *bottom-up*, gerando um mapa de saliência global. A integração de características é feita considerando a influência das saliências provenientes dos dois subsistemas.

Em relação às características definidas em ambos os subsistemas, o sistema pode ser estendido para incluir outras características baseadas inclusive em momentos que pudessem fornecer informações adicionais sobre os objetos e a cena, como simetria, desequilíbrio, alongação, cor, textura, etc. Em tarefas dedicadas é necessário um conhecimento prévio das características importantes para definir o alvo e modelar o ambiente.

Uma vez que Gestalt significa uma organização, algum trabalho envolvendo a entropia das atividades da matriz de saliência pode trazer algum resultado positivo em avaliar a conspicuidade inclusive entre cenas.

A rede *fuzzy* de duas camadas simula um paralelismo entre os subsistemas *fuzzy bottom-up* e *top-down*, entretanto ambos os sistemas podem ser considerados independentes e complementares. O subsistema *fuzzy bottom-up* pode ser utilizado para guiar o processo atencional *top-down*, permitindo por exemplo que se determine o índice de saliência *top-down* (IST) do objeto em uma tarefa específica, apenas para objeto mais conspicuo no mapa de saliência *bottom-up*. Por sua vez, o subsistema *bottom-up* leva em consideração a influência de todas as características de todos os objetos dentro da cena, entretanto a

integração pode ser realizada apenas, por exemplo, para os objetos com maior índice de *saliência top-down*. O sistema pode também incluir mecanismos de controle e recalculando o mapa de *saliência bottom-up*. Como exemplo, considere a figura 8.103 onde com relação ao alicate, todos os círculos apresentaram característica “tamanho pequeno”. Os seres humanos o tempo todo utilizam mecanismos *bottom-up* e *top-down* para interagir com o ambiente e realizar tarefas. De maneira análoga o processo *bottom-up* poderia guiar a atenção para o alicate e após verificar o seu potencial como alvo, ele poderia ser descartado e o restante da imagem poderia ser analisada novamente descartando agora sua interferência, assim o índice de atenção para o “círculo pequeno” aumentaria com a mudança de referência.

A atenção pode ser direcionada a um objeto ou a uma localização, de uma maneira *bottom-up* ou *top-down*. Em Lauwereyns (1998) pode ser vista uma discussão mostrando as relações entre essas duas dimensões da atenção, partindo inicialmente de uma pré-suposição de que essas relações referem-se a aspectos diferentes. O autor sugere inicialmente a existência de quatro tipos de seleção: controle exógeno/endógeno da atenção baseada no espaço e da atenção baseada no objeto, entretanto ele sugere que esperar um objeto não parece afetar a seleção visual e que o controle endógeno da atenção baseada no objeto pode não ser possível. O autor considera ainda que as duas dimensões da atenção visual são interdependentes, e necessitam de uma estrutura mais integrativa.

O sistema pode ser estendido de modo a permitir uma análise e detecção de regiões de interesse, variantes no tempo. As regiões extraídas pelo sistema atencional podem ser usadas para mover uma câmera ou deslocar uma janela dentro de uma imagem maior. O sistema permite também que sejam extraídas algumas informações sobre a distribuição dos objetos na cena, antes de segmentá-los, através da extração dos momentos centrais da cena global. Tais informações podem guiar a segmentação e restringir o processamento para um foco de tamanho menor. Para isso, necessita-se de um processo de controle para guardar as posições dos objetos, evitando que posições já analisadas sejam processadas novamente. Dessa forma o sistema todo pode ser utilizado para realizar a varredura de uma janela atencional que incrementalmente explora o ambiente.

A possibilidade de se restringir a busca a uma região selecionada pode reduzir o tempo de busca. Esse tipo de abordagem é interessante também pois permite selecionar regiões de processamento *acima*, *abaixo*, *à esquerda* ou *à direita*. Também é possível incluir mapas de características mais realísticos e implementar o sistema em um *hardware* paralelo e dedicado.

Existem ainda muitas maneiras de melhorar o sistema proposto e torná-lo mais adequado para aplicações específicas, como por exemplo extendê-lo para processar imagens com objetos oclusos e com fundo não uniforme pode torná-lo mais robusto e de uso mais geral.

Através do aperfeiçoamento na elaboração dos testes psicofísicos realizados para extração do conhecimento humano e aplicando-os também em um maior número de pessoas é possível de alguma maneira modelar e analisar a interferência que o ambiente exerce na realização de tarefas específicas. A validação dos resultados obtidos poderá ser realizada, por exemplo, através de outros testes psicológicos, onde o objeto selecionado pelo sistema em situações de conflito se assemelhe à tomada de decisão humana. A construção de um sistema visual que considere aspectos biológicos e psicológicos humanos ainda está longe de ser conseguida, entretanto a utilização de modelos perceptivos baseado em lógica *fuzzy* podem tornar possível a construção de modelos mais ricos que considerem os mecanismos de atenção envolvidos na percepção de formas e na realização de tarefas.

O sistema pode ser facilmente estendido para aplicações dedicadas onde a combinação complexa do relacionamento entre variáveis *fuzzy* seja necessária. Se o sistema for estendido para manipular múltiplas tarefas visuais, isto necessita de uma habilidade de decidir e controlar a seqüência de ações a serem executadas para que a tarefa corrente seja completada.

3 – CONTRIBUIÇÕES E CONCLUSÕES FINAIS

Um problema fundamental em Visão Computacional é representado pela complexidade computacional de algumas tarefas como detectar um elemento alvo em uma imagem (Busca Visual) e o problema de achar uma correspondência entre a imagem e um conjunto de modelos (*matching*). Uma solução aparente para este problema seria considerar uma certa perda da informação e estruturar o processo de *matching* em um número maior de estágios. Por exemplo, construindo-se um número de representações de propósito geral da imagem tais como orientação e mapas de características de cor, e usar estes mapas associados a informações *top-down* para acessar os modelos em uma base de dados.

Um aspecto relacionado ao uso de uma representação distribuída para múltiplos objetos é a dificuldade de integrar (*binding problem*) corretamente as diferentes características extraídas nos mapas de características. A principal causa da complexidade computacional da percepção visual consiste no aspecto espacial do problema, isto é, no fato que o número de localizações possíveis do objeto e das possíveis combinações de características é exponencial em função do número de objetos e das características. Por esta razão, um mecanismo de Atenção Visual pode reduzir significativamente a complexidade do problema, selecionando uma região de interesse na imagem na qual a correspondência com os modelos dos objetos possa ser restrita (foco de atenção).

As vantagens computacionais de tais mecanismos vêm sendo explorada para solucionar os principais problemas fundamentais na Visão aplicada. Um exemplo onde o módulo de atenção visual é de grande importância é o controle de parâmetros de uma câmera na seleção de uma região de interesse na imagem. Esta região pode corresponder tanto a objetos estáticos que atraem a atenção ou ao contrário, aos objetos móveis que estão dentro do campo de visão.

Na maioria dos trabalhos, a localização do próximo foco é computada utilizando-se um mapa retinotópico que codifica a importância de cada posição em cada estado de interpretação. Dois tipos principais de processos *bottom-up* e *top-down* contribuem para quantificar a importância de cada posição (objeto) em cada estado de interpretação. Objetos que diferem muito de seu ambiente por exemplo em cor, orientação e tamanho, atraem a atenção de uma maneira *bottom-up*. Na atenção *top-down*, os altos níveis cognitivos do cérebro influenciam o sistema atencional a selecionar em favor de uma característica em particular ou em favor da combinação de características. A primeira classe é puramente dirigida pelos dados (*data-driven*) enquanto que a segunda inclui restrições sobre a tarefa e depende do conhecimento prévio.

O sistema proposto nesta tese possuiu alguns aspectos comuns com alguns modelos atencionais como o modelo de (Koch & Ullman, 1987) que foram um dos primeiros que propuseram um modelo abstrato de atenção e incluiu conceitos de saliência e seleção de dados através de mecanismos de inibição lateral. Esses conceitos também foram as bases de outros modelos computacionais como os modelos de Ahmad, (1991), Wolfe & Cave, (1988) e Sandon (1991).

Do ponto de vista psicológico, o presente trabalho inspirou-se na teoria de percepção Gestalt para modelar o subsistema *bottom-up*. Segundo a psicologia Gestalt objetos parecidos tendem a formar um agrupamento perceptivo enquanto que objetos muito diferentes entre si ou com atributos diferentes tais como cor, brilho, contraste, textura, formato, simetria, orientação, alongação, distorção, distribuição e posicionamento dos objetos na cena (agrupamento), movimento, etc.. despertariam a atenção. Esta *diferença* (“*distância*”) entre os atributos dos objetos é chamada de *conspicuidade* (*saliência*) e significa “*aquilo que se ressalta por ser diferente*”. A conspicuidade mede quão diferente um objeto é do outro e em alguns casos pode ser útil utilizar esta abordagem para detectar rapidamente objetos alvos na imagem.

Importantes atributos da ativação *bottom-up* consideraram esses aspectos: a ativação de um item aumenta à medida que aumentam as diferenças deste item com relação aos itens vizinhos. A ativação decresce à medida que a distância entre um item e o grupo de objetos vizinhos (representados pelo centro de massa da cena) decresce também. No modelo proposto a ativação *bottom-up* seguiu ainda alguns princípios de similaridade apresentados em Duncan & Humphreys (1989) onde a similaridade entre o alvo e os distrativos aumenta, a atividade do objeto alvo diminui e a busca torna-se mais difícil. À medida que a similaridade entre os distrativos aumenta, a ativação dos distrativos diminui, e a busca se torna mais fácil, por sua vez a busca torna-se mais difícil à medida que não homogeneidade entre os elementos distrativos aumenta. A ativação *bottom-up* guia a atenção para os itens *distintos* no campo visual e não guia a atenção para um item desejado se as características deste item não forem “*incomuns*”. Por isso, a ativação *top-down*, guiada pelo usuário, é necessária. Diversos trabalhos realizados por psicólogos tentam de alguma forma medir a percepção, a atenção e interferência das características entre os processos *bottom-up* e *top-down*.¹ Através das regras de inferência da metodologia *fuzzy* apresentada aqui neste trabalho, é possível modelar muitos fenômenos perceptivos como a conjunção de características e busca assimétrica de característica entre outros.

Conforme discussão apresentada na seção 3.2 do capítulo 2, a teoria Gestalt é um sistema muito genérico de conceitos que podem ser usados para entender virtualmente qualquer aspecto de experiência e comportamento. Assim, esses princípios podem fornecer em muitos aspectos informações úteis para robôs que exibem comportamento inteligente. A psicologia Gestalt propõe que as “propriedades do todo não são o resultado da soma de suas

¹ ver Duncan & Humphreys, (1989), Wolfe, Friedman-Hill et al, (1992) e Wolfe, (1994).

partes”, assim uma parte depende do todo no qual está incluída. A dependência das partes no todo possui uma natureza variável. A psicologia Gestalt estabelece simplesmente que as partes do todo não são completamente independentes e assim não podem ser analisadas separadamente.

Muitos aspectos da teoria Gestalt têm sido utilizados em Visão Computacional. Mecanismos de atenção, uso de uma base de conhecimento e outros critérios como regularidade geométrica são raramente consideradas. Existe um aumento crescente no interesse de se incorporar aspectos perceptivos. Mecanismos de atenção, intenção e atitude são importantes no contexto de tarefas baseadas em visão ativa. Pomerantz & Pristach (1989) sugeriram que alguns aspectos da organização perceptiva como as leis Gestalt possuam importantes papéis na busca visual paralela.

Outras teoria da percepção como a Integração de características de Anne Treisman (1986) e as Rotinas Visuais de Ullman (1984) descritas no capítulo 2, também influenciaram o desenvolvimento deste trabalho.

Um grande número de estudos psicofísicos relacionados com a detecção, localização e reconhecimento de objetos no campo visual sugerem uma teoria para a percepção visual humana, em dois estágios (Kock & Ullman, 1987). O primeiro estágio consiste de um módulo “pré-atentivo”, no qual características simples são processadas rapidamente e em paralelo sobre todo o campo visual. O segundo módulo “atentivo” está relacionado com a análise de formas complexas e reconhecimento de objetos. A visão primária (*early vision*) é dita ser *bottom-up*, afetada apenas pela imagem formada na retina e pelas características neurais fixas (*hardware*). A visão superior (*higher vision*) usa os resultados provenientes de computações integrativas, expectativa, atenção, etc para processamento posterior de características salientes e comportamentalmente importantes em uma cena e para a interpretação de seus resultados.

A análise visual da forma e de suas relações espaciais exercem um papel importante no planejamento de ações e na manipulação de objetos. Esta análise não ocorre no processamento visual primário (*early processing*) mas possui grande importância nos estágios posteriores (nível intermediário - *intermediate-level*). É impossível extrair todas as propriedades relevantes de formato e todas as relações espaciais na imagem, de uma maneira *bottom-up* uniforme (*low-level*). Diferentes tarefas necessitam da extração de diferentes propriedades de formato e suas relações (Hildretch & Ullman, 1990). Segundo Ullman,

(1984) o sistema visual possui um conjunto de operações básicas que podem ser usadas para se construir diferentes rotinas visuais e para extrair uma grande variedade de propriedades de formato e suas relações. Isto implica em uma estrutura espacial e temporal do processo. Uma certa operação é aplicada em um certo ponto do campo visual e dependendo dos resultados, uma outra operação é aplicada em outro ponto.

Existem muitas razões para explicar a existência da atenção visual. O principal argumento para explicar a necessidade da Atenção Seletiva é em termos da limitação de processamento do sistema visual, entretanto, como sugerido por (Allport, 1989) a atenção pode ser explicada em um nível mais alto do que análise de característica: existe a necessidade de uma forma de seleção que assegure uma coerência comportamental (atenção para ação).

Uma vez que a percepção da visão é um meio que permite as pessoas se interagirem com os objetos que compõem o ambiente (manipular, evitar, etc), um grande número de ações são continuamente extraídas e guiadas pela percepção de objetos. Cada uma dessas ações são controladas por um grande número de parâmetros tais como identidade, posição e aparência dos objetos. O processamento seletivo também é necessário para isolar as informações que definem os parâmetros para ações apropriadas. Se cada processamento de informação gerado por algum estímulo visual, potencialmente conduz a uma ação, é necessário mantê-lo separado dos outros, entretanto eles precisam estar coordenados a fim de completar a atividade e permitir que a atenção possa ser desviada se ocorrer um novo evento importante. A atenção exerce um papel central na integração de vários atributos visuais independentes².

Neste trabalho, explorou-se o aspecto da atenção mencionado acima, onde é importante manter independentes as características (parâmetros) que controlam as ações, a fim de que diferentes combinações desses parâmetros possibilitem a realização de tarefas específicas.

Do ponto de vista biológico, o modelo apresentado aqui inspirou-se na representação de características da imagem em múltiplos mapas de características que ocorre na área visual primária (V1) do córtex visual e da combinação de características que ocorrem nos níveis mais altos do córtex, conforme relatado no apêndice 1. O processamento atencional não está

² ver Scheinerder, (1995) para uma revisão

centralizado em uma única área, mas está distribuído em uma rede de áreas, cada uma cooperando de alguma maneira para a tarefa global. As áreas corticais possuem dois papéis principais: computar os mapas de características e detectar as regiões de interesse nesses mapas.

O potencial do sistema está na capacidade de selecionar regiões na imagem baseada na presença de múltiplos canais de processamento. Os objetivos desses canais é fornecer múltiplas estratégias: para todas as imagens de interesse pelo menos um canal deve ser capaz de selecionar características salientes. Para isso assumiu-se uma representação da imagem em um conjunto de mapas independentes que extraem algumas características elementares tais como nível de cinza, tamanho, orientação, forma e distância.

Do ponto de vista computacional e psicológico o trabalho que mais se aproxima da abordagem apresentada nesta tese é a *Busca Guiada* proposta por Wolfe & Cave (1989) e Cave & Wolfe, (1990), entretanto o sistema proposto neste presente trabalho apresenta uma série de contribuições. A representação de características em um número de mapas de características, inclusive a variável distância, permitiu uma avaliação global (Gestalt) da influência do ambiente na percepção dos objetos. Através da construção das funções de pertinência dos conjuntos *fuzzy* de entrada é possível esboçar uma reação humana quanto a variáveis perceptivas de tonalidade, distância, tamanho e orientação, bem como o relacionamento entre elas. A metodologia proposta apresenta uma maior flexibilidade, controle e modelamento das situações de ambigüidade, onde os objetos “*distractors*” compartilham várias características com o objeto alvo.

O sistema proposto mostrou ter um grande potencial em ser útil em diversas aplicações, principalmente na área de inspeção automática. Além do mais o modelo apresentado abre novas perspectivas para um melhor entendimento do sistema visual humano, através da modelagem de fenômenos de alto nível, como a Percepção e Atenção Visual Seletiva. A metodologia *fuzzy* empregada aumentou o potencial do sistema pois permite a combinação diferenciada (integração) de vários canais perceptivos e independentes (*bottom-up* e *top-down*), além da extração de aspectos perceptivos humanos. Devido ao fato também da lógica *fuzzy* permitir a representação do conhecimento e inferência dos especialistas na forma de regras nebulosas que são muito próximas à linguagem natural, é fácil a extração de conhecimento através dos testes psicofísicos realizados. Inclusive é possível modelar facilmente situações onde a ordem em que é dada uma tarefa tem implicações na saída do sistema, por exemplo é possível ponderar diferentemente as regras

para as ordens “pegar o alicate escuro e inclinado” e “pegar o alicate inclinado e escuro” priorizando por exemplo a sequência semântica.

A lógica *fuzzy* simplifica o desenvolvimento, manutenção e experimentação, já que é muito mais fácil mudar regras do que programas e permite a representação de conceitos vagos, isto é, incorpora e pondera a incerteza do sistema. A lógica *fuzzy* é muito flexível e permite que vários critérios de decisão possam ser usados para unir diferentes objetivos. No modelo proposto, a maior parte do tempo computacional é consumida pelo processo de inferência, entretanto em algumas situações onde existe o conhecimento prévio sobre uma determinada faixa de característica dentro de um conjunto *fuzzy* específico, é possível definir uma *Look-up Table* para a base de regras, que reduz bastante o tempo computacional.

Optou-se por extrair tais informações de regiões segmentadas (objetos). Uma vantagem desse tipo de abordagem é que a medida de interesse está diretamente relacionada com as primitivas. Uma outra vantagem é que a região de atenção é definida para entidades macroscópicas, que revelam a presença de objetos inteiros, em oposição aos detalhes de cantos e bordas. A principal desvantagem desse método é a falta de plausibilidade biológica. O simples algoritmo de crescimento de região foi escolhido com base na simplicidade de implementação e sua eficiência computacional e não possui certamente uma equivalência biológica nos sistemas de visão humana, embora apresente plausibilidade em relação a teoria de percepção Gestalt.

Para cada região (objeto) no mapa de características, é então calculado um certo número de características independentes. Os atributos usados para “guiar” a atenção foram escolhidos com base nas evidências provenientes do estudo da atenção humana (Treisman & Gelade, 1980; Julesz, 1984): tonalidade (média de nível de cinza), tamanho, inclinação (orientação), forma e distância. Todas essas características afetam a nossa percepção e chamam a atenção de maneiras diferentes, dependendo do contexto em que se apresentam em uma cena. A característica distância usada aqui pode fornecer alguma vantagem por exemplo para os objetos mais isolados uns dos outros (Gestalt) embora neste trabalho optou-se por “facilitar” a percepção de objetos próximos ao centro de massa da cena devido também à informações de agrupamento fornecidas pelo mapa de conspicuidade de distância com relação ao centro de massa, entretanto outros critério pode ser utilizado para beneficiar outras situações.

Objetos que diferem muito dos objetos vizinhos são destacados através de uma medida de diferença entre as características, dentro de cada mapa, para cada objeto. As atividades em cada um desses mapas sinaliza quão conspícuo é um objeto em relação aos demais objetos presentes em uma cena. O grande problema com esta abordagem consiste na combinação dos diferentes mapas em uma única informação relevante para cada localização (*the binding problem*).

Para solucionar este problema, criou-se neste trabalho um subsistema *fuzzy bottom-up* cujo objetivo é combinar as características e ponderá-las de acordo com a contribuição individual de cada característica em despertar a atenção dentro do contexto global de uma cena. Um mesmo objeto pode atrair a atenção de uma maneira diferente, dependendo do contexto em que ele aparece na cena. O subsistema *fuzzy bottom-up* determina em sua saída um índice de saliência que é resultado da combinação das conspicuidades dos mapas de características de entrada. Esta combinação é determinada através de um conjunto de regras de inferência que analisam o contexto dos objetos na cena.

Um procedimento *top-down* foi adicionado para guiar a atenção para um objeto ou característica específica na cena. Por definição a atenção *top-down* está relacionada com conhecimento de alto-nível, dependente da tarefa, em termos de quais objetos devem ser localizados ou quais características devem ser consideradas para localizá-los. O subsistema *fuzzy top-down* combina então diferentes características de acordo com a relevância das mesmas na realização de diferentes tarefas, retornando um índice de saliência *top-down* em sua saída para cada objeto presente na cena. O resultado parcial para cada tarefa específica será combinado com o índice de saliência *bottom-up* no subsistema *fuzzy* de integração.

Para a implementação de um sistema que utilize múltiplos mapas de características *bottom-up* e *top-down* independentes, muitas metodologias podem ser usadas, como por exemplo redes neurais artificiais. A lógica *fuzzy* entretanto, apresenta algumas vantagens em solucionar os problemas apresentados neste trabalho, devido à sua capacidade de aprendizado, tolerância a dados imprecisos, proximidade com a linguagem natural e sua capacidade de representar variáveis subjetivas.

Comparando o procedimento de inferência de regras nebulosas com as regras convencionais, observa-se uma diferença significativa. Ao contrário das regras usuais *If...Then...*, que são usadas em lógica Booleana, as regras nebulosas podem acomodar combinações parciais.

O número de regras requeridas para uma particular inferência é bem menor do que se as regras convencionais fossem utilizadas. Os antecedentes e as conclusões de cada regra *fuzzy* são distribuição de possibilidades (Politano, 1996). Cada distribuição cobre uma certa faixa de valores do domínio. Desse modo, poucas regras são necessárias para cobrir as regiões de interesse nos domínios. Todos os valores possíveis em cada domínio não necessitam de regras separadas, como necessitam as regras convencionais. Mesmo se uma entrada em particular não combina com o antecedente da regra, sua combinação parcial com um conjunto de regras irá fornecer uma resposta. Uma outra vantagem do procedimento *fuzzy* é que ele é um procedimento paralelo. A combinação com cada antecedente da regra com um fato de entrada é computado de forma paralela. Essa característica faz o procedimento de inferência *fuzzy* apropriada para implementação em processadores paralelos. Além disso, faz o procedimento de inferência robusto à falhas e apresenta uma degradação amenizada. Como cada regra contribui (em proporção com sua combinação de entrada) para a saída final, a remoção de alguma regra (devido a qualquer falha no sistema de controle) não irá levar a sua parada. O sistema continuará operando (devido a contribuição de outras regras) apesar de não ser tão eficiente como antes. As formas exatas da distribuição dos antecedentes e consequentes não são críticas. As várias regras contribuem em paralelo quando apresentados com uma entrada particular, por isso há uma tolerância na forma das distribuições. Isto é oposto às regras convencionais onde os antecedentes e valores de saída têm de ser escolhidos cuidadosamente.

A atenção humana é resultado da interação entre diversas variáveis subjetivas, como a percepção de forma, distância, tamanho, etc. Através da metodologia empregada aqui e com a ajuda da lógica *fuzzy* foi possível extrair os graus de atenção de cada uma dessas variáveis, montar suas respectivas funções de pertinência, relacioná-las e agregá-las de modo a aplicar idéias inspiradas na percepção humana em um sistema de Visão Computacional. A metodologia proposta é simples, abrangente, flexível e permite a extração do conhecimento de diversos aspectos perceptivos humanos. Através da metodologia proposta aqui, abre-se um novo caminho para investigar a influência entre esses diversos aspectos perceptivos humanos e através dos ajustes desses parâmetros através de testes psicofísicos mais elaborados pode ser possível o desenvolvimento de uma métrica visual que estime o grau de interferência entre diversas características presentes no ambiente e as características relevantes para a execução de uma tarefa específica.

REFERÊNCIAS BIBLIOGRÁFICAS

(Abo-Zaid et al. 1988) – Abo-Zaid, A., Hinton, O R. & Horne, E. – “About moment normalization and complex moment descriptor” – in Proc. of the 4th International conference on Pattern Recognition and Image Processing, pp. 127-136, 1977.

(Abu-Mostafa & Psaltis, 1984) - Abu-Mostafa, Y.S. & Psaltis, D – “Recognitive aspects of moments invariants” – IEEE Trans. Pattern Analysis and Machine Intelligence, V.6, pp. 698-706, 1984.

(Abu-Mostafa & Psaltis, 1985) - Abu-Mostafa, Y.S. & Psaltis, D – “Image normalization by complex moments” – IEEE Trans. Pattern Analysis and Machine Intelligence, V.7, N.1, pp. 46-55, 1985.

(Ahmad, 1991) – Subutai Ahmad – “VISIT: Na efficient Computacional Model of Human Visual Attention” – PhD Thesis, University of Illinois, Urbana Champaign, 1991.

(Allport, 1989) – A. Allport – “Visual Attention “ – in M.I. Posner (edt) , Foundations of Cognitive Science – MIT Press, Bradford Books, Cambridge, 1989, pp.631-682.

(Alt, 1962) - Alt, F - “ Digital Pattern recognition by Moments” - Journal of the ACM, V.9, 1962, pp. 240-258.

(Andersen & Gnadt, 1989) – R.A. Andersen & J.W. Gnadt, - “Posterior Parietal Córtex” in Wurtz, R.H. & Goldberg M.E. (eds) – The Neurobiology of Saccadic Eye Movements – Elsevier, New York, 1989.

(Andersen & Gnadt, 1989) – R.A. Andersen & J.W. Gnadt, - “Posterior Parietal Córtex” in Wurtz, R.H. & Goldberg M.E. (eds) – The Neurobiology of Saccadic Eye Movements – Elsevier, New York, 1989.

(Andersen et al., 1985) – C.H. Andersen, G.K. Essick & R.M. Siegel – “Encoding of spatial location by posterior parietal neurons”, Science, V.230, pp. 456-458, 1985.

(Anderson & Leonard, 1958) - Anderson, N.S. & Leonard, J. A. - "The recognition, naming and reconstruction of visual figures as a function of contour redundancy", *Journal of Experimental Psychology*, V.56, 262, 1958.

(Anderson & Van Essen, 1987) – C.H. Anderson & D.C. Van Essen - "Shifter Circuits: A computational Strategy for Dynamic Aspects of Visual Processing", - *Proceedings of Nacional Academy of Science*, v.84, pp. 6297-6301, 1987.

(Anderson, 1990) – G.J. Anderson – "Focused-attention in three-dimensional space" – *Perception & Psychophysics*, V.47, pp. 112-120, 1990.

(Arbib, 1989) – "Arbib, M. A – "The metaphorical Brain 2: Neural Networks and Beyond", Wiley, New York, 1989.

(Arbib, 1993) – Arbib, M. A. – "Schemas and Neural networks: A multi-level approach to natural and artificail intelligence" – In Arbib, M. A., *Proc. Workshop on neural architectures and distributed AI:from schema assemblages to neural networks*.

(Atkinson et al, 1969) – Atkinson, J., Campbell, F.W. & Francis, M.R – "The magic number 4!0: A new look at visual numerosity judgments" – *Perception*, 5, 327-334, 1969.

(Attnaeve, 1954) - Attnaeve, F.- "Some informational aspects of visual perception" – *Psychological review*, V.61, pp. 183-193, 1954.

(Attnaeve, 1957) - Attnaeve, F. - "Physichal Determinants of the judge complexity of shape" - *Journal of Experimental Psychological*, V. 53, 1957, pp. 221-227, 1957.

(Ballard & Brown, 1982) – Ballard, D.H & Brown, C.M – "Computer Vision" - Prentice Hall – 1982.

(Basak & Pal, 1995) – J. Basak, & S.K. Pal, "The Psychological motivated Connectionist System goes Object Perception", *IEEE Transactions on Neural Networks*, V.6, N.6, November 1995.

(Baslev, 1998) – Baslev, I – "Noise tolerance of moment invariants in pattern recognition" – *Pattern Recognition Letters*, V.19, pp. 1183-1189, 1998.

- (Beck, 1966) - Beck, J. - "Perceptual grouping produced by changes in orientation and shape", *Science*, V. 154, 538, 1966.
- (Beck, 1986) - Beck, J. - "Textural segmentation", in J. Beck (ed), "Organization and Representation in Perception" - Lawrence Erlbaum Associates, hilsdale, N.J., 1982, pp. 285-315.
- (Belkasim, 1991) - Belkasim, S.O, Shridhar, M. & Ahmadi, M. - "Pattern Recognition with Moments Invariants: A comparative study and New Results" - *Pattern recognition*, V.24, N.12, pp. 1117-1138.
- (Biederman, 1973) - Biederman, I., Glass, A L., & Stacy, E.W. - "Searching for objects in real - world scenes " - *Journal of Experimental Psychology*, 97, 1, 22-27.
- (Biederman, 1987) - Biederman, I - "Recognition by Components: a theory of human image understanding" *Psychological Review*, 94, 115-147, 1987.
- (Borelli, 1999) - Borelli, J.E. - "Automação da inferência humana com o uso de lógica nebulosa: Uma Aplicação" - Escola de Engenharia de São Carlos- Dezembro de 1999 - <http://membro.intermega.com.br/celsofr/> (lógica fuzzy/ aplicação).
- (Boyce & Hossack, 1983) - Boyce, J.F & Hossack, W.J - "Moment invariants for pattern recognition" - *Pattern Recognition Letters*, V.1, pp. 451-456, 1983.
- (Brandt & Lin, 1996) - Brandt, R.D. & Lin, F. - "Representations that uniquely characterize images modulo, translation, rotation and scaling" - *Pattern recognition Letters*, 17, pp. 1001-1015, 1996
- (Bribiesca & Guzman, 1967) - Bribiesca, E. & Guzman, A - "How to describe pure form and how to measure differences in shape using shape numbers" - *Pattern Recognition*, V. 12, N.2, 1980, pp. 101-112.
- (Brown & Owen, 1967) - Brown, D. R. & Owen, D. H., - "The metrics of Visual Form" - *Psychological Bulletin*, V. 68, 1967, pp. 243-249.

- (Bruner, 1957) - Bruner, J.S. - "On perceptual readiness" - Psychological Review, V.64, 123, 1957.
- (Budrikis & Hatamian, 1984) – Budrizis, Z. & Hatamian, M. – "Moment calculations by digital filters" – ATT Bell Labs Tech. J., V.63 , 217-229, (1984).
- (Casasent & Psaltis, 1980) – Casasent, D. & Psaltis, D. – "Hybrid processor to compute invariant moments for Pattern Recognition" – Optics Letters 5, pp. 395-397, 1980.
- (Casasent et al., 1982) – Casasent, D., Cheatham, L. & Fetterly, D. – "Optical System to compute intensity moments:design" Appl. Optics, 21, pp. 3292-3298.
- (Casey, 1970) - "Moment Normalization of handprinted characters" - IBM Journal of Research and Development " - September 1970, pp. 548-557.
- (Cash & Hatamian, 1987) – G. L. Cash & M. Hatamian – "Optical Character Recognition by the method of moments" – Comput. Vision Graphics Image Process. 39, pp. 291-310, 1987.
- (Casperson, 1950) - Caspersen, R.C. "The visual discrimination of geometric forms" - Journal of Experimental Psychology, V.40, 668, 1950.
- (Cavanagh, 1990) – P. Cavanagh – "Pursuing Moving Objects with Attention", in Proceedings of the 12th Annual Conference of the Cognitive Science Society, pp.1046-1047, MIT, July, 1990.
- (Cave & Wolfe, 1990) – K.R. & J.M. Wolfe - "Modeling the Role of Parallel Processing in Visual Search", Cognitive Psychology, v.22, pp.225-271, 1990.
- (Chang et al., 1991) – Chang, C.C., Hwang, S.M. & Buehrer, J. – "A shape recognition scheme based on relative distances of feature points from the centroid" – Pattern Recognition, V.24, N.11, pp. 1053-1063, 1991.
- (Chapman, 1990) – D. Chapman – "Vision, Instruction and Action, Ph.D.Thesis, Massachusetts Institute of Technology, 1990.
- (Chen, 1990) – Chen, K. – "Efficient parallel algorithms for the computation of two-

dimensional image moments” – Pattern Recognition, V.23, N.12, pp. 109-119, 1990.

(Chen, 1993) – Chen, C-C., - “Improved Moment Invariants for shape discrimination” – Pattern Recognition, V.26, N.5, pp. 683-686, 1993.

(Cheng et al., 1998) – Cheng, H.D., Wu, C.Y., & Hung, D.L. – “VLSI for moment computation and its application to breast cancer detection” – Pattern Recognition, V. 31, N.9, pp. 1391-1406, 1998.

(Coren & Girgus, 1980) – Coren, S. & Girgus, J. S. – “Principles of perceptual organization and spatial distortion: the Gestalt illusions” – Journal of Experimental Psychology: Human Perception and performance, 6. 404-412.

(Correa, 1999) – “Correa, S.S.C - “Lógica Nebulosa” – V Escola de Redes Neurais – Conselho Nacional de Redes Neurais, pp. C073-c090, 19 de Julho de 1999 – ITA, São José dos Campos – SP.

(Costa, 1996).- Costa, J.A. F.- “Sistema de Reconhecimento de Padrões visuais invariante a transformações geométricas utilizando redes neurais artificiais de múltiplas camadas” - Dissertação de Mestrado - Escola de Engenharia de São Carlos - SP, 1995.

(Courtney et al., 1998) - Courtney, S. M. Petit, L. Maisog, J.M. Ungerleider, L. G. & Haxby, J. V. – “An area specialized for spatial working memory in human prefrontal cortex.” - Science, 279, pp.1347-1351, (1998).

(Crook, 1957) - Facsimile-generated analogues for instrumental form displays”, - In J. W. Wulfeck and J.H. Taylor (Eds.) - *Form Discrimination as related to military problems* - Washington D.C.:National Academy of Science - Nat. Res. Council, 1957.

(De Valois & De Valois, 1988) – R.L. De Valois & K.K. De Valois - “Spatial Vision”- New York, Oxford, 1988.

(Desimone et al, 1984) – R.Desimone, T.D. Albright, C.G.Gross & C.Bruce – “Stimulus-selective properties of inferior temporal neurons in the macaque – Journal of the Neuroscience, v.4, pp. 2051-2062, 1984.

(Desimone et al, 1984) – R.Desimone, T.D. Albright, C.G.Gross & C.Bruce – “Stimulus-selective properties of inferior temporal neurons in the macaque – Journal of the Neuroscience, v.4, pp. 2051-2062, 1984.

(Desimone, 1989) – Desimone, R – “Neural circuits for visual attention” – in Carpenter, G.A & Grossberg, S (eds) – Neural Networks for vision and Image Processing” – MIT Press, Cambridge, MA., 1989.

(Desimone, 1991) – Robert Desimone, “Face-Selective Cell in the Temporal Cortex of Monkeys”, - Journal of Cognitive Neuroscience, V.3, N.1, pp. 1-8, 1991.

(Deutsch, 1955) – Deutsch, J.A.A – “A theory of shape recognition” – Brit. J. Psychol., XLVI, pp. 30-37, 1955.

(Di Zenzo et al., 1992) - Di Zenzo, S., Buono, M. D., Meucci, M & Spirito, - “Optical recognition of hand-printed characters of any size position and orientation”, IBM J. Res. develop.. V. 36, N.3, May 1992, pp. 487-501.

(Didday & Arbib, 1975) – Didday, R. L. & Arbib, M. A - “Eye movements and visual perception: a two visual system model” – Man-Machine studies, 7, pp. 547-569.

(Dobnikar et al., 1992).- Dobnikar, A., Ficzek, J., Podbregar, D. e Rezar, U. - “Invariant pattern classification neural networks versus FT approach”, - Microprocessing and Microprogramming, V.33, N.3, May, pp. 161-168, 1992.

(Duda & Hart, 1973).- Duda, R. & Hart, P. E. -, “Pattern Classification and Scene Analysis” - Wiley Interscience Publication, John Wiley & Sons, USA, 1973.

(Dudani et al. 1977) – Dudani, S. A., Breeding, K. J. & McGhee, R. B., - “ Aircraft Identification by moment invariants”, - IEEE Trans. on Computers, V. C-26, N.1, January 1977, pp. 39-46.

(Duncan, 1984) – Duncan, J. – “Selectivity attention and the organization of visual information” – Experimental Psychology: General, V.113, 1984, pp. 501-517, 1984.

(Duncker, 1945) – Duncker, K. – “On problem solving” _ Psychol. Monog., V.58, 1945.

(Egeth, Virzi & Gabart, 1984) – H.E. Egeth, R.A. Virzi & H. Gabart – “Searching for Conjunctively Defined Targets”- *Journal of Experimental Psychology: Human Perception and Performance*, V.10, N.1, pp.32-39, 1984.

(Elliott, 1913) – Elliott, E.B. – “An introduction to the Algebra of Quantics”- Chelsea Publishing Company – Bronx – New York, 1913.

(Ellis, 1923) – W. D., Ellis, - “A source book of Gestalt psychology” – Harcourt Brace, Jovanovic, (1935).

(Eriksen & Hoffman, 1972) – Eriksen, C.W. & Hoffman, J. E. – “ Temporal and spatial characteristics of selective encoding from visual displays” – *Perception and Psychophysics*, 12, 201-204, 1972.

(Eriksen & James, 1986) – C.W.Eriksen & J.D. ST. James – “Visual Attention within and around the field of focal attention: A zoom lens model”- *Perception & Psychophysics*, 1986, V.10, N.1, pp. 225-240.

(Estes, 1972) – Estes, W. K. – “Interactions of signal and background variables in visual processing” – *Perception & Psychophysics*, 12(3), 278-286, 1972.

(Faber & Stokely, 1988) – Faber, T.L. & Stokely, E. M. – “Orientation of 3D Structural in medical Images” – *IEEE Trans. on Pattern Analysis and Machine Intelligence*, V.10, pp. 626-633, 1988.

(Fantz, 1961) – Fantz, R. L. – “The origin of form perception” – *Scientific American*, 204(5), 66-72, 1961.

(Farah, 1990) - Farah, M.J. – *Visual Agnosia* (MIT Press, Cambridge, MA, 1990)

(Feldman, 1985) – Four frames suffice: A provisional model of vision and space” – *Behavioral and Brain Sciences*, V. 8, pp. 265-313, 1985.

(Felleman & Van Essen, 1991) - Felleman, D.J. and D.C. Van Essen – “Distributed hierarchical processing in the primate cerebral cortex.” - *Cerebral Cortex*, 1, pp. 1-47, 1991

(Ferreira, 1988) – Ferreira, Aurélio Buarque de Holanda - “Dicionário Aurélio Básico da Língua Portuguesa” – Editora Nova Fronteira, 1988.

(Fischler & Firschein, 1987) – Fischler, M.A & Firschein, O. – “The Eye, the Brain and the Computer – Addison-Wesley Publishing Company, 1987.

(Flusser et al., 1996) – Flusser, J., Suk, T. & Saic, S. – “Recognition of blurred images by the method of moments” – IEEE Trans. On Image Processing, V.5, N.3, March 1996.

(França, 1999) – França, C.A - “Avaliação da qualidade de placas de madeira através de um sistema de inferência *fuzzy* baseado em redes adaptativas “ – Tese apresentada ao Instituto de Física de São Carlos, Universidade de São Paulo, para a obtenção do título de Doutor em Ciências: Física Aplicada, Sub-área Física Computacional. São Carlos, agosto de 1999.

(Freeman & Saleh, 1988) - Freeman, M.O & Saleh, B.E.A – “Moments invariants in the space and frequency domains” – Journal of Optical Society of America A, V.5, p. 1073

(Freeman, 1978) – Freeman, H. - “Shape description via the use of critical points” – Pattern Recognition, V.10, pp. 159-166, 1978.

(Freeman, 1980) – Freeman, H. - “Lines, curves, and the characterization of shape” - Report no. IPL-TR-80-004, Image Processing - Rensselaer polytechnic Institute, Troy, N.Y., March, 1980.

(Fukushima, 1986) – K. Fukushima, - “A Neural Network Model for Selective Attention in Visual Pattern Recognition” – Biological Cybernetics, V.55, pp.5-15 – 1986.

(Fukushima, Imagawa & Ashida, 1991) – K. Fukushima, T. Imagawa & E. Ashida, “Character recognition with Selective Attention”, in Proc.International Joint Conference on Neural Networks, Seattle, 1991.

(Garner, 1962) – Garner, W. R. - “Uncertainty and Structure as Psychology concepts” - New York: Wiley, 1962.

(Garner, 1970) – Garner, W. R. - “The stimulus in information processing” – American

Psychologist, 25, 350-358, 1970.

(Garner, 1974) – Garner, W. R. – “The processing of information and structure”, New York:Wiley, 1974.

(Gazzaniga, 1995) – Gazzaniga, M.S. – The Cognitive Neuroscience – Gazzaniga, M.S. (ed) – MIT, 1995.

(Ghent, 1956) – Ghent, L. - “Perception of overlapping and embedded figures by children of different ages” - American Journal of Psychology, V.69, 575, 1969.

(Ghorbel, 1994) - Ghorbel, F. - “ A complete invariant description for gray-level images by the harmonic analysis approach”, Pattern Recognition Letters, V.15, pp. 1043-1051, 1994.

(Gibson, 1950) - Gibson, J.J. - “The perception of the visual world”, Boston:Houghton Mifflin, 1950.

(Gibson, 1966) - Gibson, J.J. - “The senses considered as perceptual systems”, - Boston:Houghton Mifflin, 1966.

(Gilmore & Boyd, 1981) – Gilmore, J.F. & Boyd, W.W. – “Building and bridge classification by invariants moments” – Proc. SPIE V. 292, pp.256-263, 1981.

(Giuliano et al., 1961) - Giuliano, V.E., Jones, P.E., Kimball, G.E., Meyer, R.F., Stein, B.A.: Automatic Pattern Recognition by a Gestalt Method. Information and Control, Vol.4, (1961) 332-345.

(Goldman-Rakic, 1991) – P. Goldman-Rakic, - “Representations and Process in Human Prefrontal cortices”, Talk presented at the 13th Annual Conference of the Cognitive Science Society, Chicago, August, 1991.

(Goldstein, 1989) – Goldstein, E. B. – “Sensation and Perception” – Wadsworth Publishing company – 1989.

(Gollin, 1956) - Gollin, E.S. - “Some research problems for developmental psychology” - Child Development, V.27, 223, 1956.

(Gomes & Carvalho Filho, 1994) - Gomes, H. M. & Carvalho Filho, E.C.B. - "Reconhecimento off-line de assinaturas utilizando MLP - Backpropagation e Momentum: um estudo comparativo" - Anais do I Congresso Brasileiro de Redes Neurais - Itajubá - MG, 1994.

(Gonzaga & Costa, 1996) - Gonzaga, A & Costa, J.A F - "Moment invariants applied to the recognition of objects using neural networks" - Proc. of SPIE - Applications of Digital Image Processing XIX - Denver, Colorado - 1996.

(Gonzaga, 1991) - Gonzaga, A - "Tecnologia para o reconhecimento de formato de objetos tridimensionais" - Tese de doutoramento apresentada ao Instituto de Física de São Carlos, Universidade de São Paulo, 1991.

(Gonzaga, Neves & Slaets, 1998) - Gonzaga, A., Neves, E.M.A. & Slaets, A.F.F. - "Selective Visual Attention in Object Recognition and Scene Analysis", - Proceedings of SPIE's International Symposium on Optical Science, Engineering and Instrumentation - 12-24 July 1998, San Diego, Califórnia USA, pp.254-264.

(Gonzales & Wintz, 1987) - Gonzales, R. C. & Wintz, P. - "Digital Image Processing" - Addison-Wesley, Reading, Massachusetts - 1987.

(Goodale & Milner, 1992) - Goodale, M.A & Milner, A D - "Separate visual pathways for perception and action" - Trends of Neuroscience V. 15, pp. 20-25, 1992.

(Goodinatilake & Kherbal, 1995) - Goodinatilake, S., Kherbal.: Intelligent hybrid systems. John Wiley & Sons. (1995).

(Gregory, 1973) - Gregory, R.L. - "Eye and Brain" - (2nd ed.) New York, McGraw Hill, 1973.

(Gross et al, 1972) - C.G.Gross, C.E. Rocha-Miranda & D.B.Bender - "Visual Properties of neurons in inferotemporal cortex of the macaque", Journal of the Neurophysiology, v.35, pp.96-111, 1972.

(Grossberg, 199) - Grossberg, S. - "How does a brain build a cognitive code?"

(Gupta & srinath, 1987) – Gupta, L. & Srinath, M.D. – “Contour sequence moments for the classification of closed planar shapes” – pattern Recognition, V.20, N.3, pp. 267-272, 1987.

(Gurevich, 1964) – Gurevich, G.B – “Foundations of the Theory of Algebraic Invariants” – Nordoff, Groningen, The Netherlands – 1964.

(Guzman & Bribiesca, 1978) - Bribiesca, E. & Guzman, A - “Shape description and shape similarity measurement for two-dimensional regions” - “Proc. of the 4th International Joint Conference on Pattern recognition, Kyoto, Japan, Nov. 7-10, 1978, 608-612.

(Hake, 1966) – Hake, H. W. – “Form discrimination and the invariance of form” – in Pattern Recognition: Theory, Experiments, Computer simulations, and dynamic models of form, perception and discovery – L., Uhr, ed., pp. 142-173, Wiley, New york, 1966.

(Hall & Park, 1986) – Hall, E.L., & Park, K.T. – “Form recognition using moment invariants for three dimensional perspective transformation” – Science on Form: Proc. of the First International Symposium for Science on Form, S. Ishikawa (Ed.), pp. 255-264, 1986, KTK Scientific Publishers, Tokyo.

(Hall et al., 1971) - Hall, E. L., Kruger, R.P., Dwyer, S.J., Hall, D., McLaren, R.W. & Lodwick, G.S. – “A survey of preprocessing techniques for radiographic images” – IEEE Trans.on Computers, V. C-20, N.0, September 1971.

(Hall et al., 1975) - Hall, E. L., Crawford, W. O. & Roberts, F. E. - “ Computer classification of pneumoconiosis from radiographs of coal workers” - IEEE Trans. on Biomedical Engineering, V. BME - 22, N.6, November 1975, pp. 518-527.

(Hall, 1979) - Hall, E. L., - Computer Image Processing and Recognition – Academic Press – NY - 1979

(Hammond, 1966) - Hammond, K. R (Ed.) - The psychology of Egon Brunswik” - New York:Holt, Rinehart & Winston, 1966.

(Han et al., 1994) - Han, J. H., Kóczy, L.T., Lasslo, T., Timothy, P.: Fuzzy Hough transform

Pattern Recognition Letters, Vol.15, (1994) 649 –650.

(Hatamian, 1986) – Hatamian, M – “ A real time two-dimensional moment generating algorithm and its single chip implementation” – IEEE Trans. Acoust. Speech Signal Proces.” – V.34, pp. 546-533, 1986.

(Hayward, 1998) – Hayward, W.G – “Effects of outline shape in object recognition” – Journal of Experimental Psychology: Humana, Perception and Performance, 1998, V.24, N.2, 427-440.

(Head & Holmes, 1911) – Head, H. & Holmes, G. - “Sensory disturbances from cerebral lesions” – Brain, V.34, pp. 102:254, 1911.

(Hebb, 1949) - Hebb, D. O. - “The organization of behavior”, “New York:Wiley, 1949.

(Helson, 1933) – Helson, H. – “The fundamentals propositions of Gestalt Psychology” – Psychological Review, 40, 13-32.

(Hershenson, 1964) - Hershenson, M. - “Visual discrimination in the human newborn”- Journal Comp. Physiol. Psychology, V.58, 270, 1964.

(Hildretch & Ullman, 1990) - “The computational study on vision” - in Foundation of Cognitive Science, Posner, M. I. (ed) – MIT Press - pp. 581-630.

(Hochberg & McAlister, 1953) - Hochberg, J. & McAlister, E. - “A quantitative approach to figural goodness” - Journal of Experimental Psychology, V.46, 361, 1953.

(Hochberg et al., 1948) - Hochberg, J., Gleitman, H. & Macbride, P. D. - “Visual threshold as a function of simplicity of form”, - American Psychology, V.3, 341, 1948

(Hochberg, 1970) - Hochberg, J.E., “Attention, organization, and consciouness” In D. I. Mostofsky (Ed.), Attention: Contemporary theory and analysis (pp.99-124). New York:Appleton-Century-Crofts, *Threshold as a function of simplicity of form*, - American Psychology, V.3, 341, 1948.

(Hochberg, 1978) – Hochberg, J.E – “Perception” – Prentice-Hall, Englewood cliffs, New

Jersey, 1978.

(Holtzman & Gazzaniga, 1982) – Holtzman, J.D. & Gazzaniga, M. S. “Dual task interactions due exclusively to limits in processing resources” – Science, 218, 1325-1327.

(Hu, 1961) - Hu, M. K. - “ Pattern Recognition by Moments Invariants” – Proc. of the IRE, V. 49, N.9 - p. 1428, 1961

(Hu, 1962) - Hu, M. K. - “Visual Pattern Recognition by Moments Invariants” - IRE Trans. Inf. Theory, V. 8 - pp. 179-187.

(Hubel & Livingstone, 1985)- D.H. Hubel & M.S. Livingstone – “Complex-unoriented cells in a subregion of primate area 18”, - Nature, V.315, pp. 325-327, 1985.

(Hubel & Wiesel, 1962) - Hubel, D.H. & Wiesel, T. N. - “Receptive fields binocular interaction and functional architecture in the cat’s visual cortex”, - Journal of Physiology, V. 160, 106, 1962.

(Hubel & Wiesel, 1968) - Hubel, D.H. & Wiesel, T. N. - “ “Receptive fields and functional architecture of monkey striate cortex”, - Journal of Physiology, V.195, 215, 1968.

(Hyman & Hake, 1954) - Hyman, R. & Hake, H. W. - Cited by H.W. Hake - “Form discrimination and the invariance of form”, in L.Urh (eds.) - *Pattern Recognition* - New York: Wiley, 1966.

(Imai et al., 1976) – Imai, S., Ito, T. & Ito, S. – “Effect of inter-pattern transformation structures upon goodness judgements”, Psychological review, v.47, pp. 204-212, 1976.

(Ishikawa, 1978) - Ishikawa, S. - “ Geometrical Indices Characterizing Psychological goodness of Random shapes” - Proc. of the 4th IJCPR - Kyoto, Japan, Nov. 7-10, 1978, 414-416, 1978.

(Jain & Vailaya, 1998) – Jain, A.K. & Vailaya, A – “Shape-base retrieval:A case study with trademark image databases” - Pattern Recognition V.31, N.9, pp. 1369-1390, 1998.

(Jain, 1995) – Jain, R. – “Machine Vision” – MIT Press & McGraw-Hill Inc. – 1995.

(Jang & Gulley, 1995) - Jang, J.S.R. and Gulley, N. – “Fuzzy Logic Toolbox for Use With Matlab” – The Math Works Inc., 1995.

(Jay & Sparks, 1984) – M.Jay & D. Sparks, “Auditory Receptive Fields in primate superior colliculus shift with changes in eye position”, *Nature*, 309, pp.345-347, 1984.

(Johnston & Dark (1986)) - Johnston, W.O. and Dark, V.J., “Selective Attention”, *Annual Review of Psychology*, V.37, pp.43-75, 1986.

(Johnston & McClelland, 1973) – “Visual factors in word perception” – *Percept. Psychophys.*, 14(2), pp. 365-370, 1973.

(Julesz & Bergen, 1983) - B. Julesz & J.R.Bergen – “Textons, the fundamental elements in preattentive Vision and Perception of Textures”, - *Bell Systems Technical journal*, v.62, pp. 1619-1643, 1983.

(Julesz & Bergen, 1987) – B.Julesz & J.R.Bergen, - “Textons, The Fundamental Elements in Preattentive Vision and Perception of Textures, in Firschler M.A. & Firschein, O.(Eds), *Readings in Computer Vision: Issues, Problems, Principles and Paradigms*, Morgan Kaufmann, San Mateo, CA.

(Julesz, 1981) – Julesz, B. – “Figure and ground perception in briefly presented isodipole textures. In M. Kubovy & J. R. Pomerantz (Eds), *Perceptual Organization*. Hillsdale, NJ:Erlbaum, 1981.

(Kahan et al, 1987) - Kahan, S., Pavlids, T. & Baird, H. S. - “On the recognition of printed characters of any font and size” - *IEEE Trans. on Pattern Analysis and Machine Intelligence*, v. PAMI-9, N.2, March, 1987.

(Kahneman, 1973) – Kahneman, D. – “Attention and effort” – Englewood Cliffs, NJ:Erlbaum, 1973.

(Kanisa, 1979_ - Kanisa, G. – “Organization in vision” – New York:Praeger, 1979.

(Katona, 1940) – Katona, G. – “Organization & Memorizing” – New York: Colombia

University, Press, 1940.

(Khotanzad & Hong, 1990) - Khotanzad, A. & Hong, Y.H. – “Invariant Image Recognition by Zernike Moments – IEEE Trans. On Pattern Analysis and machine intelligence, V.12, N.5, May 1990.

(Khotanzad & Lu, 1990) - Khotanzad, A. & Lu, J. H. - “Object recognition using a neural network and invariant Zernike Features” - Proc. of IEEE Computer Society Conf. on Computer Vision and Pattern Recognition, pp. 200-205, 1990.

(Klir & Folger, 1988) – Klir, G.J. & Folger, T.A – “Fuzzy sets, Uncertainty and Informations” – Prentice Hall, Englewood Cliffs – New Jersey, 1988.

(Klir & Yuan, 1995) - Klir, George J. and Yuan, Bo - “Fuzzy Sets and Fuzzy Logic” - Theory and Applications, Prentice Hall PTR, New Jersey, 1995.

(Knapp, 1996) – Knapp, R.B. – “Fuzzy Sets and Pattern Recognition” – <http://www.engr.sjsu.edu/~Knapp/HCIFUZZY/fuzzy.htm>

(Kochen, 1975) – Kochen, M. – “Applications of fuzzy sets in psychology” - in *Fuzzy sets and their application to cognitive and decision process* – Zadeh, L.A, Fu, K-S., Tanaka, K. & Shimura, M. (eds) – Academic Press, 1975.

(Kock & Ullman, 1985) – C.Koch & S. Ullman – “Shifts in Selective Selection:Towards the underlying Neural Circuitry”, *Human Neurobiology*, V.4, pp.219-227, 1985.

(Köhler, 1929) - Köhler, W. - “Gestalt Psychology” - New York, Londres, 1929.

(Kokawa et al., 1975) – Kokawa, M., Nakamura, K., Oda, M. – “Experimental approach to fuzzy simulation of forgetting and inference process” – in *Fuzzy sets and their application to cognitive and decision process* – Zadeh, L.A, Fu, K-S., Tanaka, K. & Shimura, M. (eds) – Academic Press, 1975.

(Konish, 1983) – M. Konish – “Centrally Synthesized Maps of sensory space

(Kosslyn, 1995) - Kosslyn, S.M. *Image and Brain*, MIT Press, 1995.

- (Kosslyn, 1987) – “Seeing and Imaging in the cerebral hemisphere: A computational approach” – Psychological Review, 94 (2), pp. 148-175, 1987.
- (Kruger & Thompson, 1981) - Kruger, R. P. & Thompson, W. B. - “A technical for economic assessment of computer vision for industrial inspection and robotic assembly” - Proc. of IEEE, v. 69, N.12, December 1981, pp. 1524-1538.
- (La Berge, 1990) - D. LaBerge - “Thalamic and Cortical Mechanism of Attention Suggested by Recent Positron Emission Tomographic Experiments”, - Journal of Cognitive Neuroscience, V.2, N.4, pp.358-372, 1990.
- (Lambert, 1969) - Lambert, P. F. - “Designing Pattern Categorizers with extremal paradigm Information” - in Watanabe, S. (ed.), “Methodologies of Pattern recognition” Academic, New York, 1969, pp. 359-381.
- (Lee, 1990 a) - Lee, Chuen Chien - “Fuzzy Logic in Control Systems: Fuzzy Logic Controller, Part I” - IEEE trans.on Systems, Man and Cybernetics, vol.20, no. 2, March/April, pp.404-418, 1990.
- (Lee, 1990b) - Lee, Chuen Chien - “Fuzzy Logic in Control Systems: Fuzzy Logic Controller, Part II” - IEEE trans.on Systems, Man and Cybernetics, vol.20, no. 2, March/April, pp.419-435, 1990.
- (Lee, Kwak & Kwang, 1996) - Lee, Keon-Myung; Kwak, Dong-Hoon and Kwang, Hyung Lee - “Fuzzy Inference Neural Network for Fuzzy Model Tuning” - IEEE Trans.on Systems, Man and Cybernetics-Part B: Cybernetics, vol.26, no. 4, August, pp.637-645, 1996.
- (Leou & Tsai, 1987) – Leou, J-J & Tsai, W-H – “Automatic rotacional simmetry determination for shape analysis” – Pattern recognition, V.20, N.6, pp. 571-582.
- (Lettvin et al., 1959) – Lettvin, J.Y., Maturana, H.R., McCulloch, W.S. & Pitts, W.H. – “What the frog’s eye tells the frog brain” – Proc. of the IRE, V.47, N.11, pp. 1940-1951.
- (Levine et al, 1980) - Levine, M. D. - Youssef, Y. M., Noble, P. B. & Boyarsky, R., “The

quantization of blood cell motion by a method of automatic digital picture processing - IEEE Transactions on PAMI - V.2, N.5, 1980, pp. 444-450;

(Levine, 1985) - Levine, M. D. - "Vision in Man and Machine" - McGraw-Hill Publishing Company - 1985.

(Li, 1992) – Li, Y. – "Reforming the Teory of Invariant Moments for Pattern Recognition" – Pattern Recognition – V.25, N.7, pp. 723-730, 1992.

(Li, 1993) – Li, B-C – "A new computation of geometric moments" – Pattern Recognition, V.26, N.1, pp. 109-113, 1993.

(Liao & Pawlak, 1996) – Liao, S.X. & Pawlak, M. – "On Image Analysis by Moments" – Ieee Trans. On Pattern Analysis and machine Intelligence, V.18, N.3, pp. 254-266, March 1996.

(Lin & Wang, 1994) – Lin, W-G & Wang, S-S – "A note on the calculation of moments" – Pattern Recognition Letters, V.15, pp. 1065-1070.

(Lin, 1992) – Lin, J-C., - "Universal principal axes: an easy-to-construct tool useful in defining shape orientation for almost every of shape" – Pattern Recognition, V.26, N.4, pp. 485-493, 1993.

(Ling, 1941) - Ling, B. C. - "Form discrimination as learning cue in infants", - Comp. Psychol. Monog., V.17 (2), 1941.

(Livingston & Hubel, 1981) – Livingston, M. L. & Hubel, D. H - "Effects of sleep and arousal on the processing of visual information in cat", Nature, 291, 554-561, 1981.

(Livingstone & Hubel, 1987) – M.S. Livingstone & D.H. Hubel, "Psychophysical Evidence for Separete channels for perception of Form, Color, Movement and Depth", The journal of Neuroscience, V.7, N.11, 1987, pp. 3416-3468, 1987.

(Lockhead, 1966) – Lockhead, G. R. – "Effects of dimensional redundancy on visual discrimination""- Journal of Experimental Psychology, 72, 95-104.

- (Loncaric, 1998) – Loncaric, S. – “A survey of shape analysis techniques” – V.31, N.8, pp. 983-1001, 1998.
- (Lowe, 1985) – Lowe, D.G. – “Perceptual Organization and Visual Recognition” – Boston, MA: Kluwer, 1985.
- (Mach, 1914) – E, Mach, - “The analysis of sensations. Republished 1959. New York, Dover.
- (Mahoney, 1987) - J.V. Mahoney - “Image chunking: Defining Spatial Building Blocks for Scene Analysis” - MIT Artificial Intelligence Laboratory Technical Report N.980, 1987.
- (Maitra, 1979) – S, Maitra – Proc. of IEEE V.67, pp. 697-699, 1979.
- (Marr & Hildreth, 1980) – Marr, D. & Hildreth, E. – “Theory of edge detection” – Proc. Of Roy. Soc. London B204, 301-328 (1979).
- (Marr & Nishihara, 1978) – Marr, D. & Nishihara, H. K.. – “Representation and recognition of the spatial organization of three-dimensional shapes” – Proc. Of Roy. Soc. London B200, 269-294 (1978).
- (Marr, 1976) – Marr, D. – “Early processing of visual information” – Philosophical Transactions on the Royal society (London), 200, 264-294.
- (Marr, 1982) – Marr, D. – “Vision” – Freeman, São Francisco, (1982).
- (Master, 1994) - Master, T. - “Signal and Image Processing with Neural Networks” - John Wiley & Sons, Inc. - 1994.
- (Maunsell & Van Essen, 1991) - Maunsell, J.H.R & Van Essen, D.C.,- “Anatomical connections of the middle temporal visual area in the macaque monkey and their relationship to a hierarchy of cortical areas.” - J. Neurosc. 3, 2563-2586, (1991).
- (Maunsell, 1995) - Maunsell, M.R. – “The Brain’s Visual World: Representation of Visual Targets in Cerebral Cortex”- Science, 270, pp. 764-768, 1995.

(Merigan & Maunsell, 1993) - Merigan, W.H. & J.R.H. Maunsell. – “How parallel are the primate visual pathways?”- *Annu. Rev.Neurosc.* 16, 369-402, (1993).

(Merleau-Ponty, 1962) – Merleau-Ponty, M. – “Phenomenology of perception” , London:Routledge & Kegan, Paul, 1962.

(Michels & Zusne, 1965) - Michels, K.M. & Zusne, L. - “Metrics of visual form”, - *Psychological Bulletin*, V.63, 74, 1965.

(Milanese, 1993) - Milanese, R. - “Detecting Salient Regions in an Image: from Biological Evidence to Computer Implementation” - University of Geneva - Tese de Doutoramento - 1993.

(Milner, 1974) – Milner, P. M. – “A model of visual shape recognition” – *Psychological Review*, 81 (6), 521-535, 1974.

(Minsky & Papert, 1969) M. Minsky & S. Papert - “Perceptrons: Na Introduction to computacional geometry”- MIT Press, Cambridge, MA., 1969.

(Minsky, 1975) – Minsky, M – A framework for representing knowledge”, in P.H. Winston (Ed.), *The Psychology of Computer Vision*”, New York: McGraw Hill, 1975.

(Mishkin, 1983) – M. Mishkin, L.G. Ungerleider, & K. Macko, “Object vision and spatial vision: Two central pathways. “*Trends in Neuroscience*, 6, pp. 414-417, 1983.

(Miyashita & Chang, 1988) - Miyashita, Y. – “Neural correlate of visual associative long-term memory in the primate visual cortex.” - *Nature*, 335, 817-820.

(Moran & Desimone, 1985) - J.Moran & R. Desimone, “Selective Attention Gates Visual Processing in the Extrastriate CórteX”, *Science* , N. 229,pp. 782-784, March 1985.

(Moran & Desimone, 1985) - J.Moran & R. Desimone, “Selective Attention Gates Visual Processing in the Extrastriate CórteX”, *Science* , N. 229,pp. 782-784, March 1985.

(Mountcastle et al, 1975) - V.B. Mountcastle, J.C. Lynch, A. Georgopoulos & H. Sakata, “Posterior parietal association cortex of the monkey:command function for operations within

extrapersonal space”, - Journal of Neurophysiology, v.38, pp.871-907, 1975.

(Mountcastle et al, 1981) - V.B. Mountcastle, R.A. Anderson & B.C. Motter, - “The Influence of Attention Fixation upon the excitability of the light-sensitive neurons of the posterior parietal cortex” - The Journal of Neuroscience, V.1, N.11, pp.1218-1235, 1981.

(Mozer, 1988) - M.Mozer - “A connectionist Model of Selective Attention in Visual Perception”, in Proceedings of the 10th Annual Meeting of the cognitive Science Society, 1988.

(Mozer, 1991) - M.Mozer - The perception of multiple objects: A connectionist Approach. MIT Press, Cambridge, MA., 1991.

(Müller, 1896) - G. E. Müller - “Zur Psychophysik der Gesichtsempfindungen” - Zts. f. Psych. 10.

(Nagy & Tuong, 1970) - “Normalization techniques for handprinted numerals” - Communications of the ACM, V. 13, N.13, pp. 475-481, 1970.

(Navon, 1977) – Navon, D. – “Forest before trees: The precedence of global features in visual perception” – Cognitive psychology, 9, 353-383.

(Naylor, 1971) - Naylor, W. C. - “ Some studies in the interactive design of character recognition systems” - IEEE Trans. on Computers - V.20, pp. 1075-1086, 1971.

(Neisser, 1967) – Cognitive Psychology – Appleton, New York, 1967.

(Neisser, 1967) - Neisser, U. - “Cognitive Psychology” - New York: Appleton-century-Crofts, 1967.

(Neves & Gonzaga, 1996) - Neves, E. M. A. & Gonzaga, A. - A Multi-font Character Recognition Based on its Fundamental Features by Artificial Neural Networks” - Proc. of II Workshop on Cybernetic Vision - pp. 196-202 - IEEE Computer Science - 1996.

(Neves & Silva, 1998) - Neves, E. M. A. & Silva, P. C. - “Reconhecimento óptico de caracteres utilizando Invariantes de Momento e Redes Neurais Artificiais” - trabalho

apresentado no III Congresso Nacional de Iniciação Científica e III Congresso Nacional de Pesquisadores - ASSER - Novembro de 1998.

(Neves et al., 1999) - Neves, E. M. A., Gonzaga, A & Slaëts, A.F.F. - "Target Search by Bottom-up and Top-down Shape Information" - I International Seminar on Bioelectric Interfaces and III Workshop on Cybernetic Vision" – Fevereiro de 1999, pp. 109-114..

(Neves, 1995) - Neves, E. M. A. & - "Reconhecimento Óptico de Caracteres Isolados baseados em suas Características Fundamentais" - Dissertação de Mestrado - Escola de Engenharia de São Carlos - 1995.

(Newman & Jain, 1995) - Newman, T. S. & Jain, A. K. - " A survey of Automated Visual Inspection" - Computer Vision and Image Understanding, V. 61, N.2 - March, pp. 231-262.

(Norton & Stark, 1971) – Norton, D & Stark, L. – "The eye movements and visual perception" – Scientific American, 224(6), 34-43.

(O'Connel & Treisman, 1991) - K.M. O'Connell & A. Treisman - "Abstract Coding of Orientation as a Visual Feature. Perception & Psychophysics.....

(O'Regan, 1988) – J.K. O'Regan – "Visual Acuity, Lexical Structure and Eye Movements in Word Recognition", in Working Models of Human Perception, Academic Press, London, 1988, pp.261-191.

(Olshausen, Anderson & Van Essen, 1992)- B. Olshausen, C. Anderson & D. Van Essen – "A neural model of Visual Attention and Invariant Pattern Recognition", - California, Institute of Technology, Computacional and Neural System Program, CNS Memo18, August 1992.

(Otsu, 1986) – Otsu, N. – "Recognition of shape and transformation: an invariant theoretical foundation" – Sciece on Form: Proc. of the First International Symposium for Science on Form" – Ishikawa, S (Ed.), pp. 413-420, 1986, Tokyo.

(Pal & Mitra, 1992) - Pal, Sankar K. and Mitra, Sushmita - "Multilayer Perceptron, Fuzzy Sets and Classification" - IEEE Trans. On Neural Networks, vol.3, no. 5, September, pp.683-697, 1992.

- (Pal & Pal, 1993) - Pal, N. R. & Pal, S. K. - "A Review on Image Segmentation Techniques" - Pattern recognition, V. 26, N. 9, pp. 1277-1294.
- (Palmer et al., 1981) – Palmer, S., Rosch, E. & Chase, P. – "Canonical perspective and the perception of objects" – In Long, J. & Baddley, A (eds), Attention and Performance, IX, Hillsdale, New jersey: Lawrence Erlbaum, 1981
- (Papoulis, 1965) –Papoulis, A – "Probability, Random Variables and Stochastic Processes" – McGraw-Hill Kogakusha, Ltd., 1965.
- (Parker, 1994) - Parker, J. R. - "Practical Computer Vision using C" - John Wiley & Sons, USA, 1994.
- (Pei & Lin, 1995) – Pei, S-C & Lin, C-N, - "Image normalization for pattern recognition" – Image and Vision Computing, V.13, N.10, December 1995.
- (Perenin & Vigheto, 1988) - Perenin, M. & Vighetto, A - "Optic ataxia:a specific disruption in visuomotor mechanisms I. Different aspects of the deficit in reaching for objects"- Brain, 111, 643-674 (1988)
- (Petersen et al, 1985) - S.E. Petersen, D.L. Robinson & J.D. Morris - "Contributions of the pulvinar to visual spatial attention", Neuropsychology, V.25, pp. 97-105, 1987.
- (Phaf, R. H. et al, 1990) - Phaf, R.H., Van dar Heidjen, A.H.C. and Hudson, P.T.W., "SLAM: The connectionist model goes attention in visual selection tasks", Cognitive Psychology, 22, pp.273-341, 1990.
- (Piaget, 1951) - Piaget, J. - "Play, Dreams and Imitation in childhhod", - London: Heinemann, 1951.
- (Piaget, 1952) - Piaget, J. - "The origins of Intelligence in Children", - New York: International University Press - 1952.
- (Piaget, 1955) - Piaget, J. - " The child's construction of reality"- London: Routledge & Kegan Paul, 1955.

(Pink, 1984) – Pinker, S. – “Visual Cognition: na introduction” – Cognition 18, 1-63, 1984.

(Politano, 1996) – Politano, P.R. – “Proposta de um procedimento de programação reativa on-line de FMS” – Tese de Doutorado apresentada à Escola de Engenharia de São Carlos, da Universidade de São Paulo, Departamento de Engenharia Mecânica – 1996.

(Pomerantz & Kubovy, 1986) – Pomerantz, J. R. & Kubovy, 1986 – “Theoretical approaches to perceptual organization. In K. Boff, L. Kaufman & J. Thomas (Eds), Handbook of perception and human performance. New York Wiley, 1986.

(Pomerantz & Schwaitzberg, 1975) – Pomerantz, J. R. Schwaitzberg, S. D. – “Grouping by proximity: selective attention measures” – Perception and Psicophysics, 18, 355-361.

(Pomerantz, 1977) – Pomerantz, J. R. – “Pattern goodness and speed encoding” – Memory and Cognition, 5, 235-241.

(Pomerantz, 1986) – Pomerantz, J. R. – “Visual form perception: an overview” – in Pattern Recognition By Humans and Machines, V.2, Schwab, E.C. & Nusbaum, H. C. – Accademic Press, 1986.

(Portezani, 1999) – Portezani, C.H. – “Investigação e modelagem de saliências em formas bi-dimensionais” – Dissertação de Mestrado apresentada no Instituto de Física de São Carlos – USP – 1999.

(Posner & Petersen, 1990) - M.I. Posner & S.E. Peterson - “The Attention System of the Human Brain” – Annual Review of Neuroscience, v.13, pp.25-42, 1990.

(Posner et al., 1982) - M.I. Posner, Y. Cohen & R.D. Rafal - “Neural System Control of Spatial Orienting”- Phil. Trans. R. Soc. Lond., B 298, pp.187-198, 1982.

(Posner et al., 1987) - M.I. Posner, A.W. Inhoff, F.J. Friedrich & A. Cohen – “Isolating Attentional Systems: A cognitive-anatomical Analysis”, - Psychobiology, V.15, N.2, 1987, pp. 107-121.

(Posner et al., 1987) - M.I. Posner, A.W. Inhoff, F.J. Friedrich & A. Cohen – “Isolating Attentional Systems: A cognitive-anatomical Analysis”, - Psychobiology, V.15, N.2, 1987,

pp. 107-121.

(Pratt, 1960) - Pratt, K.C. - "The neonate", - In L.Carmichael (Ed.) - Manual of Child Psychology", New York: Wiley, 1960.

(Pratt, 1991) – Pratt, W.K. – "Digital Image Processing" – Jonh Wiley & Sons, Inc., 1991.

(Prokop & Reeves, 1992) - Prokop, R. J. & Reeves, A. P. - " A Survey of Moment-Based Techniques for Unoccluded Object Representation and Recognition", - CVGIP: Graphical Models and Image Processing, V. 54, N.5 - pp. 438-460, 1992.

(Rao et al., 1997) - Rao, C.S., Rainer, G. & Miller, E. K.-" Integration of what and where in the primate prefrontal cortex ." - Science, 276, pp. 821-824 (1997).

(Raveendran et al., 1993) - Raveendran, A & Yegnanarayana, B. - "A Two stage neural network for translation, rotation and size-invariant pattern recognition" - Proc. of 1991 Intl. Conf. on Acoustics, speech and Signal Processing - ICASS- IEEE V. 4, pp. 2393-2396, 1993.

(Raveendran et al., 1993) - Raveendran, P, Jegannathan, S. & Omatu, S. – "New regular moments invariants to classify elongated and contracted images" – Proc. of 1993 International Joint Conference on Neural Networks, V.3, pp.2089-2093, 1993.

(Ravichandran & Yegnanarayana, 1991) – Ravichandran, A & Yegnanarayana, B. "A two stage neural networks for translation, rotation and size-invariant visual pattern recognition" – Proc. of. 1991 Intl. Conf. On Acoustics, Speech and Signal Processing – ICASSP, IEEE, V.4, pp. 2393-2396.

(Ray & Ghoshal, 1997) – Ray, Kumar S., Ghoshal, J. – "Neuro Fuzzy Approach to Pattern Recognition" – Neural Networks, V.10, N.1, pp. 161-182, 1997.

(Rayner, 1978) – K. Rayner – "Eye movements in reading and information processing" – Psychological Bulletin, 1978, V.85, N.3, pp.618-660.

(Reddi, 1981) – Reddi, S. – "Radial and Angular Moment invariant for image classification" – IEEE Trans. On Pattern Anal. Mach. Intell., V.3, pp. 240-242, 1981.

(Reeves & Rostampour, 1981) – Reeves, A.P. & Rostampour, A - “Shape analysis of segmented objects using moments “ – in IEEE Computer society Society Conference on pattern Recognition and Image Processing, August 1981, pp. 171-174.

(Reeves & Taylor, 1989) – Reeves, A.P. & Taylor, R. – “Identification of three dimensional objects using range information” – IEEE Trans. Pattern Anal. Mach. Intell., V.11, pp. 403-410, 1989.

(Reeves & Wittner, 1983) – Reeves, A.P., & Wittner, B.S. – “Shape analysis of three dimensional shape analysis using the method of moments” – in Proc. of 1983 IEEE Conf. On Computer Vision and Pattern Recognition, June 1983, pp. 20-26.

(Reeves et al., 1988) – Reeves, A.P., Prokop, R., Andrews, S & Kuhl, F. – “Three dimensional shape analysis using moments and Fourier descriptors” – IEEE Trans. Pattern Analysis Mach. Intell. V.10, pp. 937-943, 1988.

(Reeves, 1982) – Reeves, A.P. - “A parallel Mesh moment computer “ - Proc. of the 6th International Conference on Pattern Recognition, Munich, Oct., 19-22, 1982, pp. 465-467.

(Rueckl et al, 1989) – J.G. Rueckl, K.R. Cave & S.M. Kosslyn – “Why are “What” and “Where” Processed by Separate Cortical Visual Systems? A Computational Investigation”- Journal of Cognitive Neuroscience, V.1, N.2, 1989, pp.171-186.

(Rumelhart, 1980) – Rumelhart, D. E. – “Schemata: The building blocks of cognition” – In Shapiro, R. J., Bruce, B. C. and Brewer, W. F. (eds), Theoretical Issues in Reading comprehension, Hillsdale, New Jersey: Lawrence Erlbaum.

(Sadjadi & Hall, 1978) - Sadjadi, F. A. & Hall, E. L., - “Numerical computations of moment invariants for scene analysis” – Proc. IEEE Computer Society Conference on Pattern Recognition and Image Analysis, PRIP-78, Chicago, May 31 - June 2, 1978, pp. 181-187.

(Sadjadi & Hall, 1980) - Sadjadi, F. A. & Hall, E. L., - “Three-dimensional moment invariants” – IEEE Transactions on PAMI, PRIP-78, Chicago, May 31 - June 2, 1978, pp. 181-187.

(Sakata et al., 1995) - Sakata, H., Taira, M., Murata, A. & Mine, S., "Neural-mechanisms of visual guidance of hand action in the parietal cortex of the monkey". *Cereb. Cortex* 5, 429-438, (1995).

(Salapatek & Kassen, 1966) - Salapatek, P. & Kessen, W. - "Visual scanning of triangles by human newborn"

(Sandon, 1990) - P.A. Sandon - "Simulating Visual Attention" - *Journal of Cognitive Neuroscience*, V.2, N.3, pp. 213-231.

(Sarkar & Boyer, 1993a) - Sarkar, S. & Boyer, K. L. - "Perceptual Organization in Computer Vision: A review and a proposal for a classificatory structure" - *IEEE Transactions on Systems, Man and Cybernetics*, V.23, N.2, 1993.

(Sarkar & Boyer, 1993b) - Sarkar, S. & Boyer, K. L. - "Integration, Inference and Management of spatial information using Bayesian networks: perceptual organization" - *IEEE Transactions on Pattern Analysis and Machine Intelligence*, V.15, N.3, 1993.

(Schiller & Lee, 1991) - P.H. Schiller & K. Lee- "The role of the primate extrastriate area V4 in vision", *Science*, n. 261, pp. 1251-1253, March 1991.

(Schiller, 1985) - P.H. Schiller - "A model for the generation of visually guided saccadic eye movements", in Rose, D. & Dobson, V.G. (Eds) - *Models of the visual cortex* - John Wiley & Sons, NY, 1985.

(Schiller, 1985) - Schiller, P. H. - "A model for the generation of visually guided saccadic eye movements" - "Models of the Visual Cortex" - in Rose, D. & Dobson V.G. (Eds) - John Wiley & Sons, NY, 1985.

(Sejnowski & Churchland, 1993) - Sejnowski, T.J. & Churchland, P.S., in *Foundations of Cognitive Science*, (eds Posner, I.) pp. 301-356, (MIT Press, Cambridge, MA, 1993).

(Sekuler & Blake, 1990) - R. Sekuler & B. Blaker, "Perception", McGraw-Hill, (2th edition), 1990.

(Serenio & Maunsell, 1998) - Sereno, A. & Maunsell - "Shape Selectivity in the Lateral

Intraparietal Posterior Cortex in Primates” – Nature, V.395, 1998.

(Shepard, 1964) – Shepard, R.N. – “Attention and the metric structure of stimulus space” – Journal of mathematical psychology, 1, 54-87.

(Shiffrin et al, 1976) – Shiffrin, R.M., McKay, D.P. & Shaffer, W.O – “Attending to forty-nine spatial positions at once” – Journal of Experimental Psychology, Human Perception & Performance, 2 (1), 14-22, 1976.

(Silver et al., 1966) - Silver, C.A., Landis, D. & Messick, S. - “Multidimensional analysis of visual form: an analysis of individual differences”, - American Journal of Psychology, V.79, 62, 1966.

(Simpson et al., 1991) – Simpson, M.L., Schmoyer, R.L. & Hunt, M.A – “Moments invariants for automated inspection of printed material” – Optical Engineering – April 1991, V.30, N.4.

(Slagle & Dixon , 1980) - Slagle, J. R. & Dixon, J. K. - “ Finding a good figure that approximately passes through given points” - Proceedings of the 5th International conference on Pattern recognition “ - Miami Beach, Dec. 1-4, 1980, pp. 1049-1054.

(Sluzek, 1995) – Sluzek, A. – “Identification and inspection of 2-D objects using new moment-based shape descriptors” – Pattern Recognition Letters, V.16, 687-697, 1995.

(Smith & Wright, 1971) - Smith, F. W. & Wrigth, M. H. - “Automatic ship photo interpretation by the method of moments” - IEEE Transactions on Computer, V. C-20, N.9, September 1971, pp. 1089-1095.

(Smith, 1987) – Smith, B.J. – “Perceptual Organization in a random stimulus”, in Human and machine Vision II, ^a Rosenfeld (Ed), New York: Academic , pp. 237-242.

(Sparks, 1986) - D.L. Sparks - “Translation of sensory signals into commands for control of saccadic eye movements:role of primate superior colliculus”, Physiological Reviews, v.66, n.1, 1986.

(Sparks, 1986) - D.L. Sparks - “Translation of sensory signals into commands for control of

saccadic eye movements:role of primate superior colliculus”, *Physiological Reviews*, v.66, n.1, 1986.

(Spillman & Werner, 1990) – Spillman, L. & Werner, J.S. – “Visual Perception: The Neurobiological Foundations” – Academic Press, Inc., Berkeley, CA., 1990.

(Stenson, 1966) – Stenson, H.H. – “The physical factor structure of random forms and their judge complexity” – *Perception & Psychophysics*, V.1, pp.303-310, 1966.

(Stenson, 1968) – Stenson, H.H. – “The psychophysical dimension of similarity among random shapes”, *Perception & Psychophysics*, V.3, pp.201-214, 1968.

(Strachan et al. 1990) – Strachan, N.J.C., Nesvadba, P. & Allen, A.R – “Fish species recognition by shape analysis of images” – *Pattern Recognition*, V.23, N.5, pp. 539-54, 1990.

(Sugeno & Takagi, 1985)] - Sugeno, Tomohiro and Takagi, Michio - “Fuzzy Identification of Systems and Its Applications to Modeling and Control” - *IEEE Trans. On Systems, Man and Cybernetics*, vol.SMC-15, no. 1, January/February, pp.116-132, 1985.

(Sugisaka & Teshnehlal, 1993) - “Fast pattern recognition by using moment invariants computation via artificial neural networks” - *Control - Theory and Advaced technology* - V. 9 - N.4 - December, pp. 877-886, 1993.

(Taguchi et al.Iguchi, S & Umetani, Y. – “A method of evaluation of feature parameters of convex plane figures” – *Proc. of 15th Conference of Society of Instrument and control Engineers*, pp. 447-448, 1976 (in japanese)

(Tanaka, 1996) – Representation of Visual Features of objects in the inferotemporal cortex” – *Neural Networks*, V.9, N. 8, pp.1459-1475, 1996.

(Teague, 1980) - Teague, M.R. - “Image Analysis via the general theory of moments” - *J. Opt. Soc. Am.*, - V.70, N. 8, pp. 920-930., 1980.

(Teh & Chin, 1986) – Teh, C.H. & Chin, R. T. – “On digital approximation of moment invariants” – *Computer Vision, Graphics and Image Processing*, V.33, pp. 318-326, 1986.

(Teh & Chin, 1988) – Teh, C.H. & Chin, R. T. – “On Image analysis by the methods of moments” – IEEE Trans. On Pattern Analysis and machine Intelligence, V.10, N.4, pp. 496-513, July 1988.

(Thomas, 1965) – Thomas, H. - “Visual-fixation responses of infants to stimuli of varying complexity”, Child Development, V.39, 629, 1965.

(Thorpe, 1988) - S.J. Thorpe, “Traitement d'images chez l'homme. Technique et Science Informatiques, V.7, N.6, 1988, pp.517-525.

(Titchener, 1910) – Titchener, E. B. - “Text-Book of Psychology” - New York, 1910.

(Treisman & Gelade, 1980) - A. Treisman & G. Gelade - “A Feature Integration Theory of Attention” – Cognitive Psychology, 12, pp. 97-136, 1980.

(Treisman & Gormican, 1988) - A. Treisman & S. Gormican - “Feature Analysis in Early vision: Evidence from search asymmetries” - Psychological Review, v.95, n.1, pp.15-48, 1988.

(Treisman & Sato) - A. Treisman & S. Sato - “Conjunction Search Revisited”, - Journal of Experimental Psychology: Human Perception and Performance - V.16, N.3 - pp.459-478.

(Treisman & Schmidt, 1982) - A. Treisman & H. Schmidt - “Illusory Conjunction in the perception of objects” - Cognitive Psychology, V.14 - pp. 107-141.

(Treisman, 1986) – Treisman, A. – “Features and objects in visual processing” – Scientific American, 255, pp. 114-125, 1986.

(Treisman, 1986) – Treisman, A. – “Preattentive processing in vision” – in Rosenfeld, A. (ed), Human and Machine Vision II, Academic Press, Boston, 1986.

(Treisman, 1988) - A. Treisman - “Features & Objects”, - The Fourteenth Bartlett Memorial Lecture: The Quarterly Journal of Experimental Psychology, V.40 A, n.2, 1988.

(Troxel, 1976) - Troxel, D. E. - “Feature selection for low error rate OCR “ - Pattern

Recognition, V.8, pp. 73-76, 1976.

(Tsotos, 1988) – Tsotos, J.K – “How does human vision beat the computational complexity of visual perception” – in Pylyshyn, Z.W (ed) – Computational Process in Human Vision – Ablex, Norwood, New Jersey, 1988.

(Ullman, 1984) – S. Ullman, - Visual Routines – Cognition, V.18, pp.97-159, 1984.

(Ungerleider & Haxby, (1994) - Ungerleider, L.G & Haxby, J.V. – “What and Where in the human brain” Current Opinion Neurobiol. 4, 157-165, 1994.

(Ungerleider & Mishkin, 1982) – L.G. Ungerleider & M. Mishkin – “Two Cortical Visual systems”, - in Inle, D.J., Goodale, M.A. & Mansfield, R.J.W. editors, The Analysis of Visual Behavior, Cambridge, Massachusetts, MIT Press, 1982.

(Ungerleider, 1995) - Ungerleider, L.G., Functional Brain Imaging Studies of Cortical Mechanisms for Memory, Science, 270, 769-775, 1995

(Uttal, 1975) – Uttal, WR – “Na autocorrelation Theory of form detection” – Laerence Erlbaum Associates, Hillsdale, N.J., 1975.

(Van de Laar et al, 1997) - Van der Laar, P., Heskes, T. and Gielen, S., “Task -dependent learning of attention”, Neural Networks, V.10, N.1, pp. 1-21, 1997.

(Van Essen & Anderson, 1990) - D. Van Essen & C.H. Anderson - “Information Processing Strategies and Pathways in the primate retina and visual cortex”, in: Na Introduction to neural and electronic networks - Zornetzer, S.F., Davis, J.L. & Lau, C. (eds), - Academic Press, 1990.

(Van Essen & Deyoe, 1995) - D. Van Essen & Deyoe, E A – “Concurrent Processing in the Primate Visual Cortex” – in The Cognitive Neuroscience – Gazzaniga, M.S. (ed) – MIT – 1995.

(Van Essen et al., 1991) – Van Essen, D.C. & Olshausen, C.H, Anderson, C.H. & Gallant, J.L. – “Pattern Recognition, Attention and Information Bottlenecks in the primate visual system” – Proc. SPIE Conf. on Visual Information Processin: from Neurons to Chips,

V.1473, pp.17-28, 1991.

(Verhagen, 1975) – Verhagen, C.J.D.M – “Some general remarks about Pattern Recognition: its definition; its relation with other disciplines; A literature survey” – Pattern Recognition, V.7, pp. 109-116, 1975.

(Vernon, 1974) - Vernon, M. D. - “Percepção e Experiência” - Ed. Perspectiva, 1974.

(Von der Heydt et al, 1984) – R. von der Heydt, E. Peterhans & G. Baumgartner, “Illusory Countours and Cortical Neuron Response”, Science, V. 224, pp.1260-1262, 1984.

(Weiss, 1993) – Weiss, I. – “Geometric Invariants and Object Recognition” _ International Journal of Computer Vision, V.10, N.3, pp. 207-231,1993

(Weisstein & Wong, 1986) – Weisstein, N & Wong , E. – “ Figure-ground organization and the spatial and temporal responses of the visual system” - in Pattern Recognition By Humans and Machines, V.2, Schwab, E.C. & Nusbaum, H. C. – Accademic Press, 1986.

(Wertheimer, 1912) - Wertheimer, M. - “Experimentelle Stuidien uber das Sehen von Beugung” - Zeischrift fuer Psychologie, V.61, pp. 161-265, 1912.

(Wertheimer, 1923) - Wertheimer, M. - “Untersuchungen zur lehre von der Gestalt II”- Psych. Forsch. 4, pp. 301-350, 1923.

(Wertheimer, 1925) - Wertheimer, M. - “Über Gestalttheorie”- Symposion 1, pp. 1-24, 1925.

(Wertheimer, 1945) - Wertheimer, W. - “Productive Thinking” - Harper & Row, New York, 1945.

(Wiesel & Hubel, 1963) – Wiesel, T. N. & Hubel, D. H. – “single cell responses in the stirate cortex of kittens deprived of vision in one eye”, - Journal of Neurophysiology, 2, 1003-1017, 1963.

(Wise & Desimone, 1988) - S.P. Wise & R. Desimone - “Behaviorial Neurophysiology: Insights into Seeing and Grasping”, - Science, 242, pp.736-740.

(Witkin & Tenenbaum, 1983) – Witkin, A & Tenenbaum, J. - “On the role of structure in vision” – Human and Machine Vision, J. Beck, B. Hope and A. Rosenfeld, New York: Academic, 1983, pp. 481-453.

(Wolfe, Cave & Franzel, 1989) - J.M. Wolfe, K.R. Cave & S.L. Franzel - “Guided Search: Na alternative to the modified feature integration model for visual search”, - Journal of Experimental Psychology: Human Perception and Performance, 15, pp.419-433, 1989.

(Wong & Hall, 1978) - Wong, R. Y. & Hall, E. L. - “Scene Matching with Invariant Moments” - Computer Graphics and Image Processing 8, 16-24, 1978.

(Wong et al, 1995) - Wong, W. H., Siu, W. C. & Lam, K. M. - “Generation of moments invariants and their uses for character recognition”, Pattern Recognition Letters, V. 16., N.2, pp. 115-123, 1995.

(Wong et al., 1976) - Wong, R. Y. & Hall, E. L. & Rouge, J. – “Hierarchical search for image matching” – in Proc. of 1976 IEEE Conference on Decision and Control, December 1976.

(Wood, 1996) – Wood, J. – “Invariant Pattern Recognition: A review” – Pattern Recognition, V.29, N.1, pp. 1-17, 1996.

(Wurtz & Goldberg, 1989) - Wurtz, R.H & Goldberg, M,E (eds) – The neurobiology of saccadic eye movements – Elsevier, New York, 1989.

(Yakali et al., 1998) – Yakali, H.H. Dorst, L. & Kröse, B. – “Pose characterization by independent moment-based image features of planar objects” – RWC Technical Report – Faculty of Mathematics and Computer Science – University of Amsterdam – January - 1998.

(Yazdi et al, 1993) - Yazdi, N., Ahmadi, M, & Shridhar, M. - “Multi-layer neural network architecture with external weight memory” - Proc. of the 36 th Midwest Symposium on Circuits and Systems, Vol.2, pp. 1288-1291.

(Young, 1986) – Young, R.A – “ The Gaussian Derivative Model for Machine Vision: Visual Cortex Simulation” – TR GMR-5323, GM Research Laboratories, 1986.

(Yüceer & Oflazer, 1993) - Yüceer, C. & Oflazer, K. - "A rotation , scaling and translation invariant pattern recognition classification system", Pattern recognition, V.26, N.5, May, pp. 687-710.

(Zadeh, 1965) - Zadeh, L.A. - "Fuzzy Sets" - Information and Control, vol.8, 1965, pp.338-353.

(Zadeh, 1978) - Zadeh, L.A. - "Fuzzy Sets as a basis for a theory of possibility" - Fuzzy sets and Systems, 1:3-28, 1978.

(Zakaria et al, 1987) - Zakaria, M. F. , Vroomen, L.J., Zsombor-Murray, Van Kessel, J.M.H.M. - "Fast algorithm for the computation of moment invariants" - Pattern Recognition, V.20, N.6, pp. 639-643, 1987.

(Zeki, 1978) - Zeki, S. M. - "Functional specialization in the visual cortex of the rhesus monkey" - Nature, 274, 423-428.

(Zhan & Minemoto, 1993) - Zhan, Q. & Minemoto, T. "Moment invariants for pattern recognition under translation, rotation, and changes of size & brightness" - Journal of Modern Optics, V.40, N.11, pp. 2267-2274, 1993.

(Zhou et al., 1991) - Zhou, X. & Koch, M. W. & Roberts, M. W. - "A selective attention neural network for invariant recognition of distorted objects", - Intl. Joint Conf. on Neural Networks, Publ. by IEEE, Part 2, p. 937

(Zingales, 1978) - Zingales, M. - "A organização da criatividade", - EDUSP, 1978.

(Zobrist & Thompson, 1975) - Zobrist, & Thompson, - Building a distance function for Gestal grouping" - IEEE Transactions on Computers, V. c-4, N.7, July, 1975, pp. 718-728, 1975.

(Zusne, 1970 a) - Zusne, L. - "Visual Perception of form" - Academic Press, New York, 1970.

(Zusne, 1970 b) - Zusne, L. - "Contemporary Theory of Visual form Perception III - The global theories, ch.4, pp. 108-174, " - Academic Press, New York, 1970.

UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE FÍSICA DE SÃO CARLOS

**ESTRATÉGIA ATENCIONAL PARA BUSCA VISUAL E
RECONHECIMENTO INVARIANTE DE OBJETOS
BASEADA NA INTEGRAÇÃO DE CARACTERÍSTICAS
*BOTTOM-UP E TOP-DOWN***

APÊNDICE

EVELINA MARIA DE ALMEIDA NEVES

**São Carlos
2000**

APÊNDICE 1

FISIOLOGIA DA FORMA

1 - INTRODUÇÃO

Embora pareça ser realizada de uma maneira fácil e sem esforços, a visão humana representa um problema complexo de ser resolvido. É comum a idéia de que quando vemos o cérebro faz uma cópia ou modelo do mundo externo, mas na verdade isso não acontece. O processo de visão começa nos olhos. Os olhos são apenas sensores que captam parte da informação total disponível a partir de um conjunto de imagens bidimensionais distorcidas projetadas na retina. Essas informações são então codificadas em diversas localizações do nosso cérebro.

Quanto mais se sobe na escala evolutiva, mais o processamento é deixado ao cérebro e menos aos olhos. De uma maneira simplista podemos definir o ato de ver como o armazenamento físico de um padrão de energia de luz recebido do mundo que nos cerca, enquanto que a percepção, que é um processo mais complexo, faz uma interpretação do que foi visto. Nos seres humanos o cérebro é o órgão que faz a interpretação do que é visto.

(Goldstein, 1989) define percepção como todas as experiências causadas pela estimulação dos sentidos. A importância dos sentidos foi reconhecida pelos primeiros psicólogos que raciocinaram que pensamento, aprendizagem, memória e emoção dependem da informações recebidas do mundo externo pelos nossos sentidos.

Segundo Vernon, (1974) nem toda a percepção se baseia em experiência. Ao contrário, existe um mecanismo visual primário que abrange os olhos, os nervos ópticos, os núcleos geniculados laterais e o córtex estriado, constituídos inatamente de tal forma que a discriminação de brilho, cor, movimento e formas simples, ocorre independentemente de qualquer aprendizagem através da experiência. Apesar disso, à partir da infância tais funções são cada vez mais subordinadas aos processos cognitivos de ordem mais elevada que ocorrem em outros níveis do córtex e interagem com funções de excitação da formação reticular, do tronco cerebral e do tálamo. Portanto a forma simples e a percepção do movimento são integradas e suplementadas por processos de identificação, classificação e

codificação através da operação de esquemas perceptivos que em grande parte dependem da aprendizagem, memória, atenção, raciocínio e linguagem. Indiscutivelmente os processos perceptivos simples continuam a atuar e na realidade apresentam dados sensoriais, dos quais dependem os processos mais complexos.

No caso dos olhos das rãs, por exemplo, cenas estáticas resultam em pouca quantidade de informações armazenadas ou transmitidas, enquanto que objetos móveis que podem representar alimentos ou inimigos, transmitem quantidades significativas de informação para seu cérebro. Um trabalho clássico de extração de características de uma imagem visual, com bases psicológicas foi feito por Lettvin et al (1959). Eles perceberam que os olhos das rãs utilizam quatro diferentes tipos de informações do sinal visual: (1)detetores de borda, (2)detetores de contraste móvel, (3) detetores de escuro, (4) detetores de bordas convexas.

A partir da década de 70, intensificaram-se os estudos com animais de ordens superiores e com isso o entendimento de que percepção visual é consideravelmente mais do que uma coleção de detetores neurais de características especializadas. Não há uma repartição da cena em elementos significativos e sim propriedades geométricas e relações entre objetos ficam retidas de alguma forma acessível (Fischler & Firschein, 1987).

O reconhecimento de objetos e eventos, bem como de suas características espaciais e temporais, faz parte da percepção e fornece os fundamentos da vida cognitiva do organismo, permitindo a construção de memórias de longa duração e modelos do ambiente. As tarefas fundamentais da percepção consiste em analisar o campo visual em diferentes objetos e eventos, classificar esses objetos e estabelecer suas inter-relações e agregar conhecimento e significado a eles. Tais operações são essenciais para acumularmos a base de conhecimento sobre o mundo, trocar informações, e escolher entre diferentes cursos de ações. Em resumo, sem percepção operações cognitivas seriam impossíveis.

O processo de reconhecimento refere-se à tarefa de encontrar e rotular partes de uma imagem de uma cena que correspondam aos objetos. A tarefa de reconhecer envolve dois problemas primários. Primeiro é necessário identificar os objetos e partes dele que ajudem a aplicar o conhecimento ganho anteriormente, isto é, o modelo para encontrar novos objetos (generalização). Segundo, rastrear ou localizar os objetos adequadamente (Basak & Pal, 1995).

No processo de reconhecimento em si, as memórias visuais e os mecanismos de processamento interno exercem uma grande influência em como uma cena é interpretada, as vezes até maior do que a informação fornecida pelos órgãos sensoriais externos.

O processo de segmentação que separa os objetos do fundo obviamente necessita de um mecanismo diferente daquele utilizado na rotulagem dos objetos. A interpretação simbólica é consideravelmente influenciada pela experiência adquirida. Isso evidencia que a visão humana é caracterizada por processos hierárquicos. Enquanto que o processo de segmentação é uma tarefa de baixo nível, o reconhecimento do objeto em si envolve tanto tarefas de baixo nível quanto de alto nível. Até hoje a visão de baixo nível é mais facilmente entendida do que a de alto nível. Processamento visual de baixo nível é dirigido puramente pelas propriedades dos estímulos (*data-driven*). Os processos de alto nível, por definição, são aqueles que produzem e utilizam imagens mentais e baseiam-se no armazenamento prévio de informações sobre objetos e eventos (*Knowledge-driven*). Esses processos interferem em todo o processamento cortical da imagem visual de entrada.

Um interessante e abrangente livro que envolve diversos estudos relacionados à Neurofisiologia desde o nível molecular até a experiência da consciência humana pode ser visto em Gazzaniga (1995). Uma revisão detalhada sobre os aspectos biológicos da visão e reconhecimento de objetos pode ser visto em Spillman & Werner, (1990), Wurtz & Goldberg, (1989), Levine, (1985), Van Essen & Anderson, (1990) e Schiller, (1985). No próximo capítulo serão mostrados os principais aspectos da atenção visual, bem como os principais modelos humanos e computacionais.

2- ETAPAS DO PROCESSO DE PERCEPÇÃO VISUAL

A percepção visual é um processo muito complexo. Primeiro porque raramente vemos imagens estáticas devido ao fato de que nossos olhos e nosso corpo estão quase sempre em movimento. Quando uma pessoa vê uma cadeira, por exemplo, e se move em direção a ela, seu movimento cria um movimento em sua retina, a qual envia novos sinais ao cérebro, criando assim novas percepções a serem experimentadas por ela. Essas novas percepções por sua vez, criam novas possibilidades de ação, habilitando a pessoa por exemplo, a se aproximar da cadeira e sentar nela. Goldstein, (1989) simplifica o processo de

percepção, dividindo o processo de perceber uma cadeira em sete etapas, conforme pode ser visto na figura A1.1:

- 1 - A luz atinge a cadeira e é refletida nos olhos da pessoa.
- 2 - Uma imagem da cadeira é formada na retina, uma rede de células que reveste o fundo do olho.
- 3 - Sinais elétricos são gerados pelos receptores, estruturas apropriadas para captar energia do ambiente e convertê-la em eletricidade.
- 4 - Esses sinais elétricos são transmitidos para o cérebro por neurônios, células especializadas na transmissão de eletricidade no sistema nervoso.
- 5 - Esses sinais elétricos atingem o cérebro.
- 6 - Esses sinais elétricos são processados ou analisados pelo cérebro.
- 7 - A pessoa percebe a cadeira.

Sob um ponto de vista mais preciso o processo de percepção inclui mais duas etapas:

- 8 - A pessoa reage a essa percepção, geralmente por movimento.
- 9 - Esse movimento muda os estímulos que atingem a pessoa, e o processo continua.

Figura A1.1 - Representação esquemática de alguns passos envolvidos no processo de percepção (Goldstein, 1989)

Embora seja possível estabelecer o processo de percepção em etapas sequenciais, é mais correto pensar nesse processo como ocorrendo continuamente e em alguns casos, afetando uns aos outros. Assim, os sinais elétricos que atingem o cérebro criam um momento de percepção da cadeira, novos sinais são gerados pelos receptores e enviados para o cérebro para criar um novo momento de percepção, e assim, à medida em que nos movemos em relação a ela temos uma percepção contínua.

Devemos considerar também o papel dos processos cognitivos tais como pensamento e memória, como sendo determinantes e resultantes dos processos de percepção. Nossa percepção de cadeira pode disparar pensamentos sobre a cadeira, e nossa memória de cadeiras vistas no passado, podem influenciar nossa percepção da cadeira. O processo de percepção envolve uma interação entre a informação que estimula os receptores e as informações de experiências passadas que estão armazenadas dentro de nós.

3 – ESTRUTURA E ORGANIZAÇÃO DO SISTEMA VISUAL DOS PRIMATAS

A partir da década de 80, os estudos relacionados a fisiologia da visão e da atenção visual intensificaram. Recentes progressos em Neurociência clareou a idéia da estrutura e organização dos centros cerebrais que cooperam com as funções visuais. *Especialização, modularidade e organização hierárquica* parecem ser princípios importantes para entender o funcionamento desse sistema paralelo extremamente complexo (Milanese, 1993).

O cérebro recebe informações visuais derivadas das imagens focalizadas nas retinas. As informações sensoreadas são mapeadas em diversas localizações do sistema nervoso central através de uma série de projeções paralelas. Muitos estudos neurofisiológicos do córtex visual têm sido direcionados ao entendimento de como os neurônios representam informações na imagem da retina.

A criação de representações da imagem da retina, é entretanto um pequeno componente da visão. Visão é um processo ativo que seleciona uma parte limitada da imagem visual para concentrar a atenção, isso implica na filtragem da informação irrelevante e na adição de interpretações e inferências sobre o alvo a ser identificado. Estudos em macacos têm demonstrado que este preparo dos sinais visuais começam nos primeiros estágios do córtex cerebral. O que o observador está tentando ver e o que o observador conhece sobre a cena visual tem um impacto considerável no que é representado no córtex visual. Esses estudos mostram que a maioria das áreas do córtex não ponderam igualmente todas as partes da imagem retinal (Maunsell, 1995).

Recentemente uma maior atenção tem sido direcionada ao monitoramento de representações extra-retinais¹ no córtex visual. Uma grande variedade de sinais extra-retinais tem sido descrita, incluindo mudanças na atividade neural relacionada à posição ou ao movimento dos olhos ou às mudanças de atenção a um estímulo em particular ou a uma localização espacial, mas poucos esforços têm sido direcionados a distribuição dos mesmos através de diferentes regiões do córtex visual. Existem muitas razões para acreditar que diferentes regiões do córtex visual possam ser especializadas em vários tipos de sinais extra-retinais, devido ao seu alto grau de especialização no processamento dos sinais que emergem da retina.

O córtex do macaco contém mais de 30 áreas visuais distintas (Felleman & Van Essen, 1991), conforme pode ser visto na figura A1.2². Cada uma dessas áreas possui uma representação, algumas vezes limitada, do campo visual e acredita-se exercer um papel distinto no processamento da informação visual. As áreas visuais corticais estão organizadas de uma maneira hierárquica, onde as áreas nos níveis mais altos são responsáveis pelo aumento na complexidade da informação visual. Por exemplo, enquanto neurônios da área visual primária V1 respondem fortemente a aparição de retas em direções particulares, muitos neurônios nos estágios mais altos do lóbulo temporal respondem apenas à padrões ou formas complexas, incluindo rostos e mãos (Desimone, 1991).

Para se ter uma idéia da complexidade da rede nervosa humana, basta lembrarmos que o homem dispõe de cerca de 10 bilhões de neurônios, sendo que em média cada neurônio tem sob controle outros 100 neurônios, portanto temos 1000 bilhões de conexões. No córtex cerebral a ação de um só neurônio pode influenciar até 5000 neurônios.

A performance do sistema visual humano é surpreendente. Estima-se que podemos reconhecer cerca de 100.000 objetos entre 100-200 ms (Thorpe, 1988). Além disso o reconhecimento possui alto grau de invariância à posição rotação e escala. Um dos fatores mais importantes que determina essa capacidade é o alto número de elementos processadores trabalhando no cérebro humano (aproximadamente 10^{11}). Mas esse fato sozinho não é suficiente para explicar essas performances, se considerarmos o baixo tempo

¹ Na caracterização da funcionalidade do córtex visual é útil a distinção entre sinais relacionados às propriedades físicas do estímulo da retina e sinais “extra-retinais”, provenientes de fontes diferentes da retina, que carregam informações relativas ao contexto comportamental.

² Van Essen Lab – Macaque Surface - Based Atlas <http://stp.wustl.edu/>

de resposta de cada neurônio, e se compararmos com a quantidade de dados de entrada (aproximadamente 10^8 receptores). De fato o sistema visual humano não foi estruturado para explorar o poder computacional de uma única rede de células, toda conectada, mas sim é organizado em diversas áreas especializadas na análise de diferentes aspectos da imagem. Assim, diferentes áreas representam gargalos da informação, os quais reduzem consideravelmente a quantidade de informação a ser transmitida (Van Essen et al., 1991).

Figura A1.2 – Córtex do macaco dividido em modalidades funcionais. O córtex visual é organizado em diversas áreas especializadas na análise de diferentes aspectos da imagem.

Além da disposição hierárquica, também existe segregação paralela no córtex visual. Grandes regiões do córtex visual parecem estar segregadas em dois fluxos de processamento, ou seja, os caminhos temporal (ventral) e parietal (dorsal) (Ungerleider & Mishkin, 1982). Cada um desses caminhos incluem diferentes tipos de processamento visual. O caminho parietal inclui muitas áreas que contêm neurônios com alto grau de seletividade à direção e é considerado importante para a análise de movimento e de relações espaciais (áreas V1, V2, V3, MT, MST). O caminho temporal inclui outras áreas (V1, V2, V4, e áreas TEO e TE do córtex IT) cujos neurônios parecem ser mais seletivos à padrões ou formas.

A percepção de objetos que podem ser reconhecidos em diferentes pontos de vista, em diferentes tamanhos e em diferentes posições no espaço é conseguida também em parte através da separação da informação retinal nesses dois caminhos visuais. Pesquisas realizadas por Mishkin et al, (1983) mostraram que o caminho que atinge a área parietal é crucial para localizar objetos (“ONDE”), enquanto que atinge a área temporal é crucial para o reconhecimento da forma (“O QUE”). Como há essa separação da informação, também há a necessidade de combinar essas informações a fim de se identificar e localizar o objeto ao mesmo tempo. A seguir será realizada uma descrição dos principais componentes do sistema visual humano.

4 - PRINCIPAIS COMPONENTES DO SISTEMA VISUAL HUMANO

A luz entra em nossos olhos mas antes que possamos vê-la, o sistema nervoso precisa transformar a luz em sinais elétricos e essa eletricidade precisa chegar até o cérebro e este por sua vez precisa de alguma maneira determinar o que esses sinais significam. Após ser processada na retina, a informação visual é direcionada para duas diferentes áreas do cérebro, conforme ilustram as figuras A1.2 e A1.3. Os axônios das células ganglionárias formam o nervo óptico o qual termina no Núcleo Geniculado Lateral (NGL) que faz parte do tálamo. O processo de radiação óptica continua até o córtex visual primário. Existe um segundo caminho que vai da retina diretamente para o Colículo Superior (CS). Esse caminho exerce um importante papel para controlar os movimentos dos olhos e existe também uma boa chance de que esteja envolvido na atenção encoberta, que ocorre internamente sem o movimento dos olhos.

Figura A1.3 – Principais componentes do sistema visual dos primatas.

Nos seres vivos o órgão mais importante da visão é o que faz uma interpretação, por exemplo o cérebro nos seres humanos. Alguns outros órgãos sensoriais diferentes dos olhos podem também fornecer informação visual para o órgão interpretador, como por exemplo os ouvidos do morcego, o tato de uma pessoa cega, e os sensores de calor das cobras. Uma definição mais funcional do que estrutural é mais apropriada, assim podemos considerar visão como o processo de converter a informação sensorada em conhecimento da forma, identidade ou configuração do objeto no ambientes. Memórias de experiências visuais passadas e os mecanismos de processamento internos possuem uma influência maior em como uma cena é interpretada. Ver, como definido acima, consiste em três estágios e o sistema visual humano pode ser visto como constituído de uma câmera e um computador fornecendo transdução e processamento respectivamente (Levine, 1985):

- (1) Obtenção seletiva de luz emanada do mundo exterior. Esse estágio pode ser considerado o mais simples uma vez que há uma menor quantidade de atividade neural envolvida.
- (2) projeção ou focalização desta luz em uma superfície foto-receptiva e conversão da energia de luz em um padrão de mudança química ou atividade elétrica que está relacionada de alguma maneira específica com a cena da qual a luz foi originada. Esse segundo estágio retinal que envolve tanto uma transdução sensorial quanto um processamento celular, também está localizado no olho.

- (3) Caminhos visuais. Esse último estágio consiste de um vasto número de conexões e muitos níveis complexos de processamento. Inicia-se na retina e continua em uma hierarquia de análises em direção ao cérebro.

A retina que consiste no primeiro estágio da hierarquia é um surpreendente mecanismo de redução dos dados, pois seus elementos foto-sensíveis (100 milhões de cones e 5 milhões de bastonetes) estão organizados espacialmente de uma maneira não uniforme, além do mais respondem diferentemente à luz, possuindo diferentes sensibilidade aos comprimentos de onda.

A fóvea, que é uma pequena região da retina (aproximadamente 0.4mm) apresenta a maior densidade de cones (cerca de 150.000 cones/mm²) e não possui bastonetes. Sua alta resolução permite explorar precisamente a região central da imagem enquanto mantém uma representação vaga da periferia (sua densidade decresce radialmente em função da excentricidade). A principal desvantagem desse tipo de estratégia de sensoramento é a necessidade do movimento dos olhos que desloca a fóvea sequencialmente para as partes mais importantes da imagem. Além dos mecanismos anatômicos para a seleção da informação, existem também os funcionais, como por exemplo a *atenção* que pode ser definida como a capacidade de selecionar uma parte da imagem na retina, para a qual a aplicação de processos de níveis mais altos possam ser restritamente aplicados (Milanese, 1993). Assim a *atenção* é usada para reduzir a quantidade de informação a ser manipulada por níveis cognitivos mais altos no cérebro, enquanto mantém alta performance (Van der Laar, 1997).

A informação obtida pelos foto-receptores é então processada através de diversas camadas de células retinais, principalmente representadas pelas células horizontal, bipolar e ganglionárias. Existem dois tipos principais de células ganglionárias na retina, conhecidas como células tipo P (parvocelular) e M (magnocelular). As células P possui campos receptivos³ pequenos, permitindo processamento detalhado de características, incluindo cor. Respondem apenas à estímulos estacionários. Estão mais concentradas na retina. As células M possuem grandes campos receptivos e não são muito seletivas à processamento espacial

³ A definição clássica de campo receptivo de uma célula é a região do campo sensorial na qual um estímulo sensorial adequado irá produzir uma reação. As propriedades dos campos receptivos nos níveis mais altos do córtex visual diferem dos primeiros estágios, onde os campos receptivos são muito pequenos (1/6 ° na região da área V1) e as células são “ajustadas” para responder à estímulos simples (barras e pontos de luz). No córtex temporal inferior, por exemplo, os campos receptivos são muito maiores (10° de todo o campo visual) e o estímulo produz reações mais complexas (mãos, faces).

nem à comprimento de onda. Respondem melhor a padrões modulados na frequência de 20 Hz ou mais e estão mais concentradas na periferia. Essa distinção entre P e M é mantida quando elas atingem o NGL e o córtex. Como elas podem se inibir mutuamente, elas são vistas a originar dois caminhos visuais no córtex que são os fluxos magnocelular e parvocelular.

A saída da retina é o nervo óptico (que contém cerca de 1.2 a 1.5 milhões de axônios das células ganglionárias) e atinge dois centros do cérebro: o núcleo geniculado lateral (NGL), o qual faz parte do tálamo, e o colículo superior (CS). O primeiro caminho chamado retino-geniculado é mais estudado do que o colicular, uma vez que ele atinge o córtex onde 90% da informação visual é processada. O caminho colicular é particularmente importante para a atenção visual e movimento dos olhos.

O tálamo é um corpo relativamente grande, está localizado no centro do cérebro organizado em um número de núcleos, alguns deles estão diretamente conectados ao córtex. O tálamo pode ser pensado como uma estação de chaveamento, pois é nele que os neurônios de todos os sentidos, exceto para o olfato, são agrupados juntos em destino ao cérebro. O núcleo mais estudado é o NGL, o qual representa um estágio intermediário entre a retina e o córtex visual. O NGL está organizado em 6 camadas e é um importante dispositivo chaveador usado para segregar os canais P e M e alinhar as entradas provenientes dos dois olhos. As células LGN possuem uma correspondência de 1:1 em relação as células ganglionárias e estão organizadas retinotopicamente, ou seja respeitam a topologia da retina. A sua função exata ainda não está clara mas há hipóteses de que esteja envolvido diretamente com atenção visual. As fibras nervosas que deixam o NGL atingem então a área receptora visual no lóbulo occipital do córtex cerebral. A área receptora visual também é chamada de córtex estriado, conforme ilustra a figura A1.4 a seguir.

4.1 – O córtex visual

Não se conhece muito sobre a organização funcional detalhada do cérebro, que contém muitas outras camadas horizontais distintas de células de várias formas e tamanhos mas sabe-se que o processamento visual nos organismos vivos é composto de uma hierarquia de análises. No início da hierarquia a informação é representada de maneira retinotópica, à medida que se sobe na hierarquia o tamanho dos campos receptivos

aumentam e a sua topologia é perdida. Embora exista uma realimentação massiva entre as áreas, a organização hierárquica ajuda a entender seus papéis e também a entender todo o sistema como capaz de realizar fluxos de processamento concorrentes, responsáveis por extrair informações de cor, forma disparidade binocular e movimento.

A área occipital do cérebro, onde os caminhos sensoriais terminam é chamada de córtex cerebral. Ela contém cerca de 70% de todos os neurônios do sistema nervoso central. O córtex visual, isto é, a parte do córtex que responde aos estímulos visuais, é um complexo de várias áreas, principalmente distribuídas no lóbulo occipital, mas está também presente em porções dos lobulos temporal, parietal e frontal.

Figura A1.4 – Áreas do córtex Visual.

Todo o córtex é uma folha fina de neurônios que envolve o tálamo e está organizado em camadas. A partir do lóbulo occipital parece haver uma divisão do fluxo de informação visual em dois caminhos, entre os lobulos temporal e parietal.

As fibras nervosas de cada olho que alcançam o córtex estriado, preservam a topologia e muito da geometria da informação da imagem da cena; uma porção do córtex estriado, chamada de *área de projeção visual*, possui uma correspondência espacial de um para um com a retina. Lesões nessa área podem causar pontos cegos no campo visual consistentes com o mapeamento retina - córtex. As fibras nervosas da metade esquerda de cada retina envia informação sobre a metade direita do campo visual ao córtex estriado no lóbulo occipital do cérebro. Similarmente, a metade direita de cada retina manda informação sobre a metade esquerda do campo visual para o córtex estriado direito. Surpreendentemente

em algumas pessoas lesões parciais nessa área deixaram intacta a habilidade de reconhecimento de padrões.

Nos seres humanos, a região imediatamente vizinha do córtex estriado é chamada de *área de associação visual*, que pode ser considerada a área principal do córtex. Essa área contém cerca de 10^5 neurônios por mm^2 de superfície dispostos em camadas e compreende as informações do córtex visual primário (área 17) projetadas no córtex adjacente pré-estriado (áreas 18 e 19). Lesões nessa área do córtex occipital podem causar distúrbios na percepção de complexos visuais completos, inabilidade de combinar impressões individuais em padrões completos e inabilidade de reconhecer objetos complexos ou suas representações pitorescas.

O córtex de associação projeta-se para o córtex Temporal Inferior (TI), que corresponde às áreas 20 e 21. As áreas 17, 18 e 19 são consideradas geralmente como “córtex visual”, e são às vezes referidas como áreas visuais I, II e III (ou V1, V2, V3 respectivamente) e estão relacionadas relativamente com processamentos de baixo-nível. Essas áreas possuem um papel importante para as funções de memória associativa com padrões visuais. Em contraste, o córtex temporal inferior tem um papel significativo na aprendizagem de discriminação visual e está provavelmente envolvido em processamento visual de alto nível. A percepção obviamente necessita da atuação de todos esses centros.

A integração de percepção com funções cognitivas em seres humanos é demonstrada em pacientes com lesões nas regiões parietais e occipitais do cérebro (regiões localizadas entre a visão primária e os centros da fala). Tais pacientes possuem grandes dificuldades em realizar tarefas que envolvam relacionamentos espaciais. A percepção de figuras é afetada por lesões em qualquer lugar do córtex, mas são mais pronunciadas quando as lesões ocorrem nas áreas associadas com a fala (lóbulos temporal e parietal esquerdos). Lesões no lóbulo temporal direito interferem no entendimento de material pictórico complexo.

Não se conhece muito sobre a organização funcional detalhada do cérebro, que contém muitas outras camadas horizontais distintas de células de várias formas e tamanhos. Muito pouco pode ser dito sobre o relacionamento entre a arquitetura cerebral e a realização de funções de alto nível, mas conhece-se um pouco sobre o processamento neural mais elementar. Existem alguns truques de processamento que a natureza descobriu e que são tão importantes que são empregados em quase todos os sistemas visuais de todas as espécies.

Uma dessas habilidades é chamada de inibição lateral e é usada para detectar quando alguma diferente ou não usual ocorre no campo visual.

Cada uma das áreas NGL, V1 e V2 formam um mapa topográfico da atividade retinal, embora os detalhes variam de área para área. As camadas são conectadas em uma maneira aproximadamente hierárquica, com o tamanho dos campos receptivos aumentando à medida que se sobe na hierarquia. Essa hierarquia não é estritamente para frente uma vez que existem conexões que pulam para outras camadas por exemplo do V1 para o temporal médio (TM). As conexões para outras áreas corticais tendem a ser bidirecionais e topográficas. Os alvos de cada neurônios possuem projeções para trás para o mesmo neurônio. No NGL estima-se que o número de fibras que retornam é cerca de dez vezes o número de fibras que saem a partir dele. Cada área cortical também envia um número significativo de projeções recíprocas para o pulvinar.

Mishkin & Ungerleider (1982) foram os primeiros a sugerir que as projeções do córtex para as áreas extra - estriadas no córtex do macaco poderiam estar subdivididas em dois fluxos de processamento funcionalmente diferentes. Com base na anatomia, junto com evidências eletrofisiológicas e estudos comportamentais em macacos, Mishkin & Ungerleider sugeriram que o fluxo ventral exerce um importante papel na identificação dos objetos enquanto que o fluxo dorsal é responsável pela localização dos objetos no espaço visual. Devido ao fato do PP (córtex parietal posterior) estar envolvido na localização dos objetos, bem como na determinação de suas relações espaciais e devido ao fato do IT (córtex temporal inferior) realizar essa integração e síntese das propriedades do objeto e fornecer uma representação central do estímulo para identificação e memorização, Mishkin & Ungerleider, (1982) sugeriram que esses dois fluxos de processamento realizavam diferentes tarefas visuais: reconhecimento do objeto (O QUE), localização do objeto (ONDE). Também foram chamados de caminhos ventral e dorsal respectivamente devido a sua localização na parte ventral e dorsal do cérebro. Essa separação foi sugerida originalmente por estudo de lesões em macacos.

- Lesões no TI causaram deficiências na performance de tarefas de discriminação de padrões mas não alteraram tarefas de visão espacial.
- Lesões no córtex parietal posterior não impediu discriminação visual mas causou sérias deficiências em tarefas que necessitavam de avaliação das posições absolutas ou relativas dos objetos.

A figura A1.5 descreve a organização do córtex visual do macaco e representa uma vista lateral do hemisfério cerebral, mostrando algumas áreas visuais selecionadas. Algumas dessas áreas encontram-se totalmente encobertas em dobras corticais e não são visíveis nesta superfície de visão. O diagrama em blocos mostra as relações funcionais entre as áreas corticais. Essas áreas são determinadas em níveis, em uma hierarquia baseada nas características anatómicas definidas por suas conexões axônicas (Maunsell & Van Essen, 1991). As áreas nos níveis superiores representam estágios posteriores de processamento cortical. As áreas se dividem em dois fluxos de processamento, os quais acredita-se mediar diferentes classes de funções visuais, onde:

- AIT: (*anterior inferotemporal area*) - área inferior temporal anterior
- CIT: (*central inferotemporal area*) - área inferior temporal central
- LIP: (*lateral intraparietal area*) - área intraparietal lateral
- MT: (*middle temporal area*) - área temporal média
- MST: (*medial superior temporal area*) - área temporal superior média
- PIT: (*posterior inferotemporal area*) - área inferior temporal superior
- VIP: (*ventral intraparietal area*) - área ventral intraparietal
- V1: área visual 1
- V2: área visual 2
- V4: área visual 4
- 7^a: área visual de Broadmann

Figura A1.5 – Organização do córtex visual do macaco (Maunsell & Van Essen, 1991).

Em geral, existe uma organização hierárquica no processo que flui entre as áreas corticais. No começo da hierarquia a informação é representada de maneira retinotópica e à medida que alcançam as áreas mais altas, o tamanho dos campos receptivos aumentam e a topologia é perdida. Embora exista um grande número de comunicações de retorno entre as áreas, a organização hierárquica permite descrever o sistema todo como um processo concorrente, responsável pela extração de informações sobre cor, forma, disparidade binocular e movimento.

4.1.1 – Área Visual Primária (V1)

A área V1, também chamada de córtex estriado é a mais estudada uma vez que é a base da hierarquia. A informação que alcança o córtex divide-se em dois fluxos principais. Um é o fluxo magno, originado das células ganglionária magno-celulares, as quais são altamente seletivas à direção e está associada com análise de movimento e o outro fluxo é originado das células parvo-celulares ⁴. As células V1 podem ser classificadas como simples, complexa e hipercomplexa. As células simples são caracterizadas por campos receptivos excitatório e inibitório, que pode ser equivalentemente descrito pelas funções de Gabor, diferenças de ajuste de gaussianas (Differences Of Offset Gaussians – DOOG), ou derivadas das gaussianas (Young, 1986). As células complexas são sensíveis à análise de textura (orientação) mas insensíveis a posição e as células hipercomplexa são muito seletivas a extremidades, curvaturas e ângulos. Em resumo, pode-se pensar que V1 representa um número de filtros de propósito geral que transforma a imagem de entrada em número de *mapas de características* retinotópicos (Milanese, 1993).

Um arranjo de células simples de diferentes camadas do tecido cortical forma uma coluna e células em múltiplas colunas, as quais cobrem toda a faixa de orientações, formam uma hipercoluna. As células complexas apresenta uma forte não-linearidade que faz com que elas sejam sensíveis à orientação mas insensíveis à posição. Isso sugere que elas podem resumir a informação espalhada sobre orientação e pode ser úteis na segmentação da imagem e da análise de textura (Van Essen & Anderson, 1990). As células hipercomplexas são muito sensíveis às extremidades das linhas, curvatura e ângulos.

⁴ Na verdade são dois fluxos, parvo-blob e parvo interblob, o primeiro responsável por processamento de cor e o outro representa a principal fonte para orientação e informação de disparidade (ver Van Essen & DeYoe, 1995).

A tabela A1.1 extraída de Goldstein, (1989) resume as propriedades dos principais propriedades desses tipos de células no sistema visual. O que pode se concluir é que à medida em que nos distanciamos da retina em direção ao cérebro, mais se necessita que um estímulo seja específico para que uma célula dispare.

Tabela A1.1 - Propriedades das principais células do sistema visual (Goldstein, 1989)

CÉLULAS	CARACTERÍSTICAS
Fibra do nervo óptico (Células ganglionárias)	Campo receptivo tipo <i>centro-vizinhança</i> . Responde melhor à pequenos pontos mas também responde à outros estímulos.
Geniculado lateral	Campo receptivo tipo <i>centro-vizinhança</i> . Muito similar ao campo receptivo das fibras ópticas nervosas.
Cortical simples	Áreas “ligadas” e “desligadas” dispostas lado a lado. Respondem melhor à barras com uma orientação particular.
Cortical complexa	Não respondem à pequenos pontos. Respondem melhor à movimentos de barras orientadas em uma dada orientação sobre o campo receptivo. Muitas células respondem à direções particulares de movimento
Cortical hipercomplexa	Não respondem à pequenos pontos. Respondem à curvas, ângulos ou barras de um comprimento particular movendo-se em uma direção particular

Essa propriedade em responder apenas à estímulos específicos é mais pronunciada em alguns neurônios encontrados fora da área receptora visual, como por exemplo no lóbulo temporal do macaco foram encontrados neurônios que respondem melhor a faces do que à outros estímulos (Desimone et al, 1984).

As células V1 estão conectadas a área V2 e a outras áreas corticais. Neurônios sensíveis a orientação do estímulo, bem como as células seletivas à frequência espacial, velocidade, disparidade binocular, cor, comprimentos, extremidades, curvatura, textura e movimento foram achados nas áreas V1 e V2 (Van Essen & Anderson, 1990), mas o

conjunto completo de características ainda está por ser determinado. A área V1 contém cerca de 1000 vezes mais neurônios do que o NGL. Uma vez que existe um mapeamento aproximadamente 1:1 do NGL com as células ganglionárias, existem ainda *centenas* de mapas de características que podem ser encontrados na área V1 (Ahmad, 1991).

A atenção pode exercer sua influência realçando a atividade nos primeiros níveis do processamento visual. Evidências fisiológicas do aumento da atividade neural no córtex visual primário (V1) foi citada por Desimone et al, (1994) entretanto ainda existem controvérsias sobre o a que nível a atenção influencia a percepção. Existe entretanto um consenso geral que não é possível entender o que é percepção sem entender a memorável capacidade de alterar o processamento visual baseado na saliência e novidade dos itens ou até mesmo em uma escolha excêntrica.

4.1.2 – Áreas Visuais V2 e V3

Além do processamento da informação de cor e movimento, um dos papéis da área V2 é combinar a informação de contraste computada em V1. Isso resulta em células que detectam contornos extensos incluindo os contornos ilusórios, ou seja, contornos que são percebidos como um todo, mas que estão fisicamente interrompidos (Von der Heydt et al, 1984).

A área V3 recebe entradas de V1 e V2 e as projeta na área V4 e no temporal médio (TM), porém usas propriedades específicas não são ainda conhecidas. Hubel & Livingstone, (1985) observaram células que detectam certas características independentemente da posição em que aparecem. Essas células parecem fazer conexões OU com as células retinotópicas de características e disparam toda vez que qualquer célula retinotópica detecta uma característica. Baseado em resultados psicológicos experimentais, Treisman, (1986) e Treisman & Gelade, (1980) propuseram que cada mapa de características contém uma célula similar chamada célula de resposta combinada (*pooled response cell*). Essas células indicam qual característica está presente em qualquer lugar do mapa mas não sua posição.

4.1.3 – Área V4

Essa área está associada à constância da cor, à síntese de formas complexas e à combinações de cor e forma. Nessa área a organização retinotópica começa a diminuir. As células dessa área podem possuir grandes campos receptivos e podem abranger múltiplos objetos. A Atenção parece exercer um importante papel na redução de interferências.

Existem fortes evidências de que a atenção espacial começa a exercer um papel maior nessas áreas. relataram a evidência de que a localização e tamanho efetivos dos campos receptivos dos neurônios na área V4 e TI podem mudar de acordo com a tarefa que o animal tenta realizar (Moran & Desimone, 1985). Os campos receptivos das células na área V4 que abrangem um estímulo atendido não respondem a um estímulo não atendido. Eles mostraram que a resposta normal a um estímulo atendido indica que a atenção não intensifica a atividade, mas tende a suprimir as regiões não atendidas. Entretanto, as células respondem a um estímulo não atendido se o estímulo atendido estiver fora do campo receptivo. Células do TI mostraram características similares, exceto que elas nunca respondem aos estímulos não atendidos.

Schiller & Lee, (1991) realizaram estudos com macacos relativos aos efeitos de lesões na área V4 na capacidade visual e concluíram que a área V4 está envolvida na detecção de estímulos menos salientes, uma tarefa que necessita atenção.

4.1.4 – O Córtex Temporal Inferior (TI)

As células do TI respondem à características mais complexas tais como várias formas e texturas e até mesmo mãos e faces (Desimone et al., 1984). Elas possuem campos receptivos maiores e parecem responder à características indiferentemente de suas localizações.

Esta área recebe a principal projeção da área V4 e representa a área visual final para reconhecimento de objeto. Duas áreas distintas anatomicamente e funcionalmente podem ser descritas:

- A porção inferior (área **TEO**) a qual parece ser responsável pela construção de associações complexas das características sensoradas.
- A porção posterior (área **TE**) onde supõe-se ocorrer discriminações de alto nível.

No córtex TI não há organização retinotópica e o tamanho do campo receptivo da área TE pode cobrir todo o campo visual, tornando o reconhecimento independente da posição do objeto na retina. Muitos neurônios no TI respondem seletivamente às características dos objetos e alguns deles são especializados na combinação de características particulares, as quais cooperam na definição de objetos particulares tais como faces e mãos. O TI faz uma integração e uma síntese das propriedades do objeto e fornece uma representação central do estímulo para identificação e memorização.

Tanaka (1996) em seu trabalho sobre a representação de características do objetos no córtex TI afirma que as características extraídas na área V1 não são suficientes para representar os objetos mas também uma combinação (holística ou local) das características dos objetos é necessária.

4.1.5 – Área Temporal Média (TM)

A área **TM** (*MT - Middle Temporal*) está organizada retinotopicamente e representa uma importante rota da informação visual que vai para o córtex parietal. Esta área recebe a maior parte das entradas provenientes do fluxo M e é claramente responsável pela análise global de movimentos.

Essa área está organizada retinotopicamente e representa a principal rota da informação visual para o córtex parietal. Essa área recebe a maioria das entradas do fluxo M e é responsável pela análise global de movimento. O TM projeta-se também em outras áreas que estão envolvidas diretamente com a investigação feita pelo movimento dos olhos (rastreamento) como o TMS, e extraem os aspectos ainda mais globais do movimento, tais como a expansão e contração de um objeto que se move radialmente com respeito ao observador.

4.1.6 – Córtex Parietal posterior (PP)

Essa área do córtex recebe projeções das áreas V1, V2 e V4 , em particular do campo periférico e também do Colículo Superior (CS) que será abordado na seção seguinte. Realmente o progresso em demonstrar a seletividade de resposta das células nesta região teve que esperar o desenvolvimento da técnica *single-unit*⁵ em macacos acordados. Registros

em macacos anestesiados praticamente não mostravam atividade consistente nestas regiões. Em outras palavras estava claro que a menos que os macacos pudessem mover os seus olhos ou suas mãos, não muito acontecia.

O principal papel atribuído ao córtex PP é a localização e configuração espacial dos objetos. Para isso ele necessita manter na memória a localização dos objetos analisados previamente. O PP também é particularmente importante para a atenção e movimento dos olhos, para isso a informação sobre o alvo na retina é combinada com a posição dos olhos a fim de se obter a coordenadas. Os movimentos dos olhos podem ser voluntários ou espontâneos. Os movimentos voluntários são aqueles envolvidos em rastreamento (*tracking*) de objetos ou em tarefas espaciais por exemplo do tipo “determinar o segundo objeto à esquerda” e também estão envolvidos na investigação e extração de características da cena e localização espacial dos objetos.

A resposta de uma célula PP depende da localização do objeto e da posição dos olhos (Andersen et al., 1985). Ela dispara mais fortemente quando na posição preferida é visto da direção preferida (posição do olho). Dessa maneira ela codifica a posição do estímulo relativa ao observador. As células PP também indicam a posição relativa do objeto focalizado com respeito aos outros objetos (Ungerleider & Mishkin, 1982). O Córtex PP é particularmente importante para a atenção e para o movimento dos olhos. Foram encontradas três classes diferentes de células:

- as que respondem durante o rastreamento de estímulos móveis
- as que descarregam antes do movimento dos olhos ou atencional
- as que estão envolvidas na tarefa de fixação de objetos estacionários

A característica dessas tarefas é ser voluntária e de alguma maneira dependente da informação *topdown*. Isso é consistente com a necessidade de movimentos dos olhos

⁵ Técnica utilizada para se medir através de microeletrodos, a atividade neuronal, produzida pela ação de algum estímulo visual específico. Enquanto que as tecnologias para explorar a organização funcional do cérebro humano têm avançado rapidamente nos últimos anos, os modelos animais permanecem a única fonte de informação detalhada sobre como os neurônios codificam a informação visual. O macaco é um excelente modelo para o sistema visual humano, possuindo capacidades visuais que são comparáveis às humanas. Imagens cerebrais fornecidas por PET (do fluxo sanguíneo) e *functional magnetic resonance imaging* FMRI (da oxigenação do sangue) apresentam ainda baixas resoluções temporal e espacial comparadas às obtidas com o monitoramento da atividade neural fisiológica (Técnica Single Unit), entretanto oferecem uma maneira de estudar o cérebro humano em funcionamento e preencher as lacunas em relação ao estudo com animais (Ungerleider, 1995). Uma descrição dos principais métodos para se estudar o cérebro dos seres humanos e de animais pode ser visto em (Kosslyn, 1995).

voluntários (Andersen & Gnadt, 1989) a fim de se analisar as relações espaciais entre os objetos.

Mountcastle et al, (1981) testaram o comportamento dos neurônios sensíveis à luz no córtex PP em macacos e perceberam que a atividade nesses neurônios aumentaram quando os animais estavam fixando a atenção e propuseram que o PP incorporavam um centro de comando geral para respostas óculo - motoras. Mountcastle et al, (1975) descobriram também células sensíveis à direção dos movimentos voluntários dos olhos mas não aos espontâneos. Uma possibilidade é que essas áreas codificam um mapa de prioridades de alto nível e que o CS é responsável pelas trocas espontâneas de atenção (Ahmad, 1991).

Segundo (Goodale & Milner, 1992) as áreas PP estão envolvidas na codificação de memória de curta duração em tarefas de localizações espaciais.

(Mountcastle et al., 1975) mostrou que subconjuntos de células no córtex PP estão envolvidas em:

- Fixação visual
- Perseguição e movimento dos olhos (sacadas)
- Alcance guiado visualmente
- Manipulação de objetos

(Sakata et al, 1995) mostraram que algumas células no córtex PP (próxima a área LIP) que dispararam quando o macaco manipulou um objeto, mostraram-se também sensíveis às características estruturais do objeto como tamanho e orientação, necessárias para determinar a postura da mão e dedos durante o movimento de agarrar (*grasping*). As fontes de informação de forma não foram localizadas, mas não parecem ser provenientes do córtex temporal uma vez que macacos com lesões na área TI apresentaram deficiências no reconhecimento de objetos mas foram capazes de alcançar e agarrar objetos como os animais normais.

4.2 – Outras áreas relacionadas com a fisiologia da Atenção Visual

Além das áreas corticais descritas nas seções anteriores, outros centros visuais podem ser encontrados no cérebro, de particular interesse com a atenção e o movimentos dos olhos.

4.2.1 – Colículo Superior (CS)

A área CS também desempenha um importante papel na atenção e no movimento dos olhos, e está estruturada em sete camadas agrupadas em duas classes: superficial e profunda.

As camadas superficiais recebem entradas diretamente da retina e indiretamente e em menor proporção do córtex visual. As entradas retinais são principalmente provenientes das células ganglionárias M e formam mapas retinotópicos. Com respeito ao córtex, não é realizado um processamento complexo e hierárquico nas camadas superficiais uma vez que suas células possuem campos receptivos muito grandes e possuem pouca sensibilidade à orientação e à cores, entretanto essas células possuem propriedades temporais muito importantes, respondendo a estímulos que se movem de $0.5^\circ/\text{s}$ a $30^\circ/\text{s}$ sendo que a maioria delas não são seletivas direcionalmente. Uma importante propriedade dessas células é que elas aumentam sua atividade antes da geração de um movimento dos olhos. Uma vez que o tempo de disparo dessas células em resposta à um estímulo é menor que o de áreas semelhantes no córtex, uma função possível para o CS é de ser um módulo rápido de advertência (Milanese, 1993). Dessa maneira, ele poderia localizar estímulos visuais de importância comportamental tais como objetos contrastados ou que rapidamente aparecem no campo visual.

A maioria das células das camadas superficiais se projetam para o Pulvinar e indiretamente para as outras áreas corticais envolvidas nas funções visuais.

Ao contrário das camadas superficiais, as camadas profundas do CS recebem entradas das diversas modalidades sensoriais (visual, auditiva e somático-sensorial) e das áreas motoras e existem mapas para cada uma dessas modalidades. As entradas visuais diretas incluem o córtex parietal, o lóbulo temporal e o lóbulo occipital, mas a maior parte das entradas são provenientes do Tálamo, especificamente do NGL (Sparks, 1986),

entretanto direta ou indiretamente, as camadas profundas recebem entradas de todas as áreas visuais do córtex.

Acredita-se que o CS esteja envolvido na integração da informação de localização dessas diferentes modalidades e no comportamento geral de orientação. Isso é sustentado pelo fato de que neurônios vizinhos em diferentes mapas sempre se referem a mesma localização espacial (Jay & Sparks, 1984).

As camadas profundas do CS contém mapas de erros para o movimento de olhos (Sparks, 1986). Esses neurônios estão dispostos retinotopicamente em *clusters* mas em cada localização a atividade do *cluster* corresponde a uma coordenada motora, codificando-se assim a direção e amplitude dos movimentos dos olhos.

O CS está principalmente envolvido coma geração dos movimentos dos olhos mas possui também uma relação com atenção. Uma vez que a atenção quase sempre é mudada para uma localização antes de um movimentos dos olhos (Posner & Petersen, 1990), é possível que o mesmo *hardware* neural sirva para mapas básicos de prioridades para ambos processos.

4.2.2 – Campos Frontais dos Olhos (CFO)

Essa área recebe entradas do córtex visual e está diretamente conectada aos centros ocular-motores. Existem evidências de que os campos frontais dos olhos estão envolvidos na geração dos movimentos dos olhos e imagina-se que tenha um papel análogo ao do CS e que seja particularmente sensível a estímulos coloridos de alta resolução (Schiller, 1985) embora seu papel na atenção não seja bem entendido.

4.2.3 – O Córtex Pré-frontal (PF)

Devido a teoria prevalecente da estrutura modular do córtex e também devido ao fato do córtex pré-frontal (PF) receber entradas de ambos os fluxos temporal e parietal, trabalhos recentes têm sido focalizados no papel do córtex PF em entender aonde e como a informação sobre a identidade do objeto é integrada com (ou permanece separada de) a informação sobre a localização do objeto (Rao et al. 1997; Courtney et al., 1998).

Também existem evidências de que o córtex pré-frontal esteja envolvido com o movimento dos olhos. Boch and Goldberg, (1987) sugeriram que o córtex pré-frontal está envolvido com comportamento direcionado ao alvo e que algumas células anteciparam a aparência do estímulo.

Goldman-Rakic, (1991) argumentou que o córtex PF se comporta como uma memória de trabalho temporária para tarefas espaciais.

Em macacos, a memória de trabalho tem sido tipicamente estudada em tarefas de comparação com a amostra ou monitorando-se o tempo de atraso das respostas. Em ambos os tipos de tarefas é dado ao macaco uma pista, a qual é mantida na memória durante um período de atraso de vários segundos. No final deste período o macaco é requisitado a fazer uma escolha ou dar uma resposta diferenciada baseada na pista que foi fornecida previamente. Muitos estudos têm descoberto neurônios cuja resposta à pista inicial é mantida no período de atraso. Segundo Ungerlaider (1995) dependendo do tipo de pista, neurônios com tais propriedades têm sido encontrados no córtex temporal inferior (para pistas de cor ou padrões visuais), no córtex parietal posterior (pistas espaciais) e no córtex pré-frontal (todos os tipos de pistas).

5 – DISCUSSÃO SOBRE OS CAMINHOS VISUAIS PARA IDENTIFICAÇÃO E LOCALIZAÇÃO DE OBJETOS

Conforme mostrado nas seções anteriores, no sistema visual primário, considera-se que os estímulos que atingem a retina são processados em dois caminhos separados: *percepção do objeto e percepção espacial* (Ungerleider & Mishkin, 1982), (Mishkin et al., 1983) e (Van Essen & Anderson, 1990).

O caminho da percepção do objeto (O QUE) identifica o objeto ou característica que o animal vê. Ele começa na retina, passa pelo NLG, áreas V1, V2 e V4 do córtex visual e termina no complexo temporal Inferior (TI). As áreas V1, V2 e possivelmente V4 consistem de vários mapas retinotópicos que são sensíveis às características tais como cor, contraste, orientação, frequência espacial, bordas e linhas (De Valois & De Valois, 1988) e (Van Essen & Anderson, 1990). Quando uma característica, por exemplo uma borda orientada

está presente em uma posição particular da retina, a célula sensível a orientação na posição correspondente no mapa dispara mais ativamente.

O outro caminho da percepção espacial começa na retina, passa através do NGL, V1, V2 e a área Temporal Média (TM) e termina no complexo Parietal Posterior (PP) (Van Essen & Anderson, 1990). O caminho Onde determina a posição do objeto que o animal vê.

Com base em evidências neurofisiológicas e em outras linhas de evidências, Ungerleider & Mishkin (1982) propuseram que o caminho temporal seria importante para determinar O QUE é o objeto, codificando características tais como forma e cor, enquanto que o caminho parietal estaria mais envolvido em determinar ONDE está o objeto, codificando características espaciais dos estímulos visuais. Esta dicotomia tem mostrado ser uma valiosa estrutura para estimar as funções do córtex visual, embora a segregação funcional e anatômica desses caminhos não esteja completa (Merigan & Maunsell, 1993). Para uma *revisão* ver (Ungerleider & Haxby, 1994)⁶. Estudos neuropsicológicos em pacientes com danos em ambas as regiões temporal e parietal sustentaram essa distinção funcional (Goodale & Milner, 1992; Farah, 1990, Perenin & Vigheto, 1988). Os dois caminhos visuais também têm sido descritos em termos do tipo de tarefa comportamental que cada um sustenta. Assim, o caminho temporal é visto estar envolvido em memória visual de curta duração para forma (Miyashita & Chang, 1988) e reconhecimento de objetos, enquanto que o caminho parietal exerce um papel importante na orientação espacial e nos movimentos guiados visualmente e isso envolve memória de curta duração para localização espacial (Farah, 1990). As áreas MT e MST que compreendem o córtex parietal são freqüentemente referenciados como o caminho do movimento (*motion pathway*) devido a grande quantidade de neurônios seletivos à direção (Maunsell & Van Essen, 1991).

O monitoramento de neurônios em macacos tem sido consistente com seus supostos papéis funcionais. Por exemplo, neurônios no caminho parietal respondem diferentemente dependendo da orientação dos olhos em suas órbitas (Andersen & Mountcastle, 1991), enquanto alguns neurônios no caminho temporal são ativados seletivamente quando o animal é requisitado a relembrar um padrão particular (Sejnowski & Churchland, 1993).

⁶ No trabalho de Ungerleider & Haxby (1994) a distinção ventral-dorsal foi sustentada através da medida de fluxo sanguíneo em regiões cerebrais em humanos usando *Positron Emission Tomography* PET enquanto realizavam tarefas de identificação e localização de objetos.

Muitas investigações dos sinais extra-retinais tem utilizado tarefas adaptadas das prováveis funções da área sendo estudada e têm sido feitas algumas comparações entre os caminhos visuais usando o mesmo estímulo visual ou a mesma tarefa comportamental. Como resultado, não está claro que representações extra-retinais são especializadas em regiões particulares do córtex, entre os caminhos parietal e temporal ou entre diferentes níveis de processamento dentro de cada caminho.

Uma outra dicotomia relacionada com os dois caminhos (O QUE e ONDE) é em termos da visão focal e ambiente (ou periférica) (Sekuler & Blake, 1990). Embora os neurônios do TI possuam campos receptivos muito grandes, eles são muito afetados pelos estímulos vindos da região central da retina, enquanto os neurônios da área PP são afetados tanto pela fóvea quanto pela periferia.

Rueckl et al, (1989) fez uma análise computacional das vantagens obtidas pela divisão do sistema visual em módulos separados para tarefas de reconhecimento e localização de objetos.

Como o (reconhecimento) geralmente gera uma ação, Goodale & Milner, (1992) reinterpretaram a divisão de trabalho entre os dois caminhos ventral e dorsal em termos de percepção x ação e sugeriram que melhor que caminho visual “WHERE”, o caminho parietal deveria ser referenciado como “HOW”.

O caminho ventral é responsável pela percepção dos objetos enquanto que o caminho dorsal atua como mediador nas transformações sensoriais motoras necessárias para guiar ações diretamente àqueles objetos. À medida que nos deslocamos em direção a um objeto, informações por exemplo de orientação, tamanho e movimento referentes ao objeto mudam e o sistema visual-motor precisa constantemente recalcular estas informações de maneira *on-line* para que seja possível o alcance e manipulação precisos dos objetos enquanto que o sistema de reconhecimento do objeto permanece *invariante*. É importante perceber que ambos os caminhos processam as mesmas informações do objeto quanto a orientação, tamanho, movimento, etc., só que a natureza de processamento é que é diferente (*on-line* x *off-line*).

Uma explicação para as diferenças de processamento entre percepção e ação é a necessidade de gerarmos representações de longa duração, onde a codificação visual deve

deixar a identidade do objeto independente (invariante) de qualquer condição particular de visão. Como consequência a percepção humana é caracterizada por constância de forma, tamanho, cor, brilho e localização. A geração dessas representações baseadas no objeto depende de uma rede de múltiplas vistas a fim de que a identidade do objeto possa ser acessada por interpolação.

As transformações visuais motoras que sustentam ações em direção aos objetos devem ser centradas em relação ao observador. Em outras palavras, em outras palavras, a localização do objeto, sua orientação e movimento devem ser codificados em relação ao observador. Aqui acontece a constância de tamanho, pois o tamanho real do objeto deve ser independente de sua distancia em relação ao observador, entretanto a posição e orientação do objeto alvo raramente é constante, tais computações devem ser recalculadas toda a hora que uma ação ocorre (em tempo real).

Em resumo, a informação visual sobre forma, tamanho, orientação e localização é utilizada em ambos os caminhos embora os algoritmos de transformação aplicados a essas entradas são dependentes da função de cada sistema. Como conclusão, devemos perceber que estas distinções são apenas uma aproximação, uma vez que existem muitas ligações entre várias áreas dos dois caminhos visuais, conforme ilustra a figura A1.6 a seguir extraída de Van Essen & Deyoe (1995). Além disso, as áreas temporal inferior e parietal estão também indiretamente ligadas através do sulco temporal superior.

Recente trabalho de eletrofisiologia em macacos realizado por Sereno & Maunsell (1998) teve o objetivo de mostrar que a área LIP (que faz parte do córtex PP), apresentou seletividade à formas bidimensionais durante a realização de uma tarefa simples de fixação, quando o macaco não tinha a intenção de realizar um movimento dos olhos nem a intenção de alcançar e manipular os objetos. Outro objetivo do artigo foi examinar o papel do córtex PP em realizar uma tarefa de atenção visual, requisitando-se ao animal guardar na memória (STM) a forma do objeto (tarefa de comparação com a amostra). Nessa tarefa requisitou-se ao animal fazer uma *sacada* para a forma teste correspondente à amostra entre um conjunto de amostras. Esse estudo chamou a atenção para a interpretação convencional da representação de informação que o animal está relembando ou atendendo. Uma maior ênfase tem sido dada ao córtex PP em atuar como memória de curta duração para posição real e relativa dos objetos, entretanto nesse trabalho os macacos mostraram seletividade à forma bidimensional também na memória de curta duração. A seletividade à forma no

caminho ONDE é uma fronteira relativamente inexplorada, que pode mudar o entendimento atual do processamento cortical. As descobertas obtidas entretanto sustentam a idéia de que *à nível eletrofisiológico não existem unidades que representam o conceito do objeto mas o conceito do objeto está apenas representado pelas atividades elétricas dos neurónios distribuídos nas várias regiões do cérebro.*

Figura A1.6 – Hierarquia das áreas do Córtex dos Primatas.

Tanaka (1996), também reforçou esse conceito da distribuição dos diversos aspectos atribuídos aos objetos, relatando que a imagem visual do objeto é processada no caminho visual ventral e as representação que mostram similaridade em relação às outras representações e as associações de imagens do mesmo objeto sob diferentes condições também são reconstruídas nessa região. Informações emocionais relacionadas aos objetos são estimuladas por entradas provenientes do córtex TI são obtidas na amígdala, associações com outros objetos são obtidas no córtex *perihinal* e comportamento significativo emerge do córtex Pré-frontal. A imagem visual do objeto é também processada no caminho dorsal, e a informação necessária para manipulação dos objeto é obtida no córtex parietal.

6 - CONCLUSÕES SOBRE A ORGANIZAÇÃO DO SISTEMA ATENCIONAL

Uma primeira conclusão que se pode tirar sobre a organização do sistema atencional é que ele não está centralizado em uma única área, mas está distribuído em uma rede de áreas, cada uma cooperando de alguma maneira para a tarefa global. Uma outra conclusão importante é que as áreas envolvidas no sistema atencional também estão envolvidas no controle do movimento dos olhos e parece que a única diferença entre essas duas funções é que os olhos necessitam de um sinal de “permissão” para mover-se ao novo alvo (Milanese, 1993). Esses sinais seriam gerados pelo CS e avaliados pelo CFO, isso é compatível com o conhecimento de que o córtex frontal está associado ao papel de decisão central.

A idéia de que o sistema atencional envolve várias áreas no cérebro foi sugerida por (Posner et al, 1987). Segundo ele, a atenção seqüencialmente envolveria as seguintes etapas: primeiro seria necessário o “desligamento” do foco corrente, depois “mudá-lo” para o novo foco e então “fixá-lo” ao novo alvo. Posner propôs que o córtex PP seria responsável pelo processo de “desligamento” e a operação de “mover” o foco seria em particular do CS e o possível candidato para o processo de “fixar” o foco seria o pulvinar (núcleo que faz parte do tálamo).

O processamento atencional é originado em várias áreas do sistema visual, incluindo o córtex estriado, a maior parte do córtex extra-estriado e o CS. As áreas corticais possuem dois papéis principais: computar os mapas de características e detectar as localizações de “interesse” nesses mapas. Essa última operação pode ser realizada através de mecanismos de competição, possivelmente implementados com conexões lateralmente inibitórias. Devido

ao seu papel de avaliação espacial entre objetos, o PP seria capaz de selecionar uma região que está espacialmente relacionada com o foco atencional anterior ou seria capaz de especificar uma região particular de acordo com a tarefa (por exemplo, o segundo objeto à esquerda). O papel do CS é detectar os alvos candidatos que principalmente causam mudanças temporais repentinas devido à entrada do objeto no campo visual. Em algum estágio, deve haver uma integração desses múltiplos requerimentos de atenção. Evidências de convergência sugerem que isso pode ser realizado ao nível do pulvinar onde existem mapas retinotópicos que recebem projeções de múltiplos mapas de características corticais. O CS parece estar envolvido em um processo de integração, de maneira análoga ao pulvinar. Isso é consistente com o fato de que vários mapas retinotópicos existem em suas camadas profundas os quais resultam de múltiplas projeções das áreas corticais.

No nível cortical, o pulvinar e o núcleo reticular (ambos no tálamo) executam uma operação de “fixação”. O efeito dessa operação é facilitar o processamento dentro do novo foco de atenção. A fim de que essa operação possa ser realizada, é necessário que a área PP “desligue” o alvo anterior. Isso é importante porque as ligações diretas entre PP e IT e V4 poderiam “trancar” as operações de reconhecimento do alvo anterior (Milanese, 1993).

Em seu modelo de atenção visual humana, Ahmad, (1991) desenvolveu uma arquitetura que baseia-se com dados psicofísicos conhecidos. Em seu trabalho, a retina, o NLG e as áreas V1, V2 se encarregam de extrair características simples, o pulvinar atua como um sistema de chaveamento (*gated system*) filtrando a informação dentro do foco atencional e as áreas CS, CFO, PP atuam como mapas de prioridade *bottom-up*, sendo controlado pelo PP. O córtex PF atua como memória de trabalho e as áreas V4, TI, MT se relacionam também com mapas de características “filtrados” (*gated feature maps*).

7 – O PROBLEMA “MENTE-CORPO” (“*MIND-BODY PROBLEM*”)

Como os processos físicos que ocorrem no corpo são transformados em experiência subjetiva? Os processos físicos que conduzem a experiência subjetiva é a atividade elétrica em neurônios isolados e em grandes populações de neurônios. A nível molecular essa atividade elétrica (por exemplo percepção) é causada pelo fluxo de íons de sódio Na^+ e potássio K^+ através da membrana nervosa (bombeamento).

Podemos pensar nos sinais elétricos como um código sensorial que leva várias propriedades (qualidades) do ambiente para o cérebro, por exemplo vermelho, azul, verde, ou amarelo para a visão e salgado, amargo, doce, azedo para o paladar. O Estudo do código sensorial objetiva identificar as informações carregadas pelos disparos dos neurônios que o cérebro utiliza para a percepção. O código sensorial representa as localizações dos sinais elétricos que atingem o cérebro. A teoria da especificidade estabelece que diferentes qualidades dentro de um mesmo sentido são sinalizadas ao cérebro através da atividade em fibras nervosas específicas (apenas para aquelas qualidades). A teoria dos padrões através das *fibras* (*across-fiber-pattern theory*) estabelece que o código sensorial que diferenciam qualidades dentro de um sentido são sinalizados ao cérebro por um padrão de atividade através de um grande número de neurônios

A resposta elétrica obtida nos eletrodos (técnica *Single Unit*) não informa como ou até mesmo se a informação que está sendo utilizada pelo cérebro. Por exemplo, a gravação da atividade elétrica por eletrodos mostrou que o disparo de células simples no córtex visual contém informações sobre orientações das linhas, mas não se sabe como a informação é utilizada pelo cérebro.

Em um âmbito ainda mais geral, não se sabe como 10 bilhões de células nervosas interagem para produzir uma consciência única. A compreensão da percepção humana de um estímulo (som, sabor, etc.) em termos de impulsos elétricos que atingem o cérebro é um dos grandes problemas sem solução. Ainda não se sabe como nossa percepção de calor e frio, cores, sons, odores, sabores podem ser explicadas por diferentes impulsos elétricos.

Além de tentar responder o *mind-body problem*, o estudo da percepção tem sido motivado pela curiosidade em se conhecer também o que os animais percebem, Pesquisas têm mostrado que o macaco *rhesus* possui uma percepção comparável à humana. Outro assunto de interesse é o desenvolvimento da percepção em crianças. Recentes pesquisas afirma que muitas habilidades perceptivas desenvolvidas em crianças no primeiro ano de idade, são muito próximas às adultas (Goldstein, 1989). O estudo da percepção pode auxiliar também deficientes visuais.

APÊNDICE 2

FORMA COMPUTACIONAL BASEADA EM MOMENTO

1 - INTRODUÇÃO

O estudo da Forma é uma importante parte do campo de Reconhecimento de Padrões e Visão computacional. “Visão Computacional é a área da Inteligência Artificial envolvida na extração de informações de imagens, com a intenção de identificar e classificar objetos e padrões nelas contidos (2D ou 3D), obtendo representações de níveis mais altos para o uso do homem ou de máquinas” (Costa, 1996).

“Visão Computacional é a automatização e a integração de uma grande quantidade de processos e representações usadas para a percepção visual e envolve técnicas de Processamento de Imagem (transformação, codificação e transmissão) e classificação de padrões” (Ballard & Brown, 1982).

A área de Visão Computacional teve um enorme progresso em pesquisa e aplicações práticas nos últimos anos. Sistemas que possuem esta tecnologia estão sendo utilizados em diversas áreas como a robótica, automação industrial, sensoramento remoto, aplicações militares e em produtos comerciais. Muitas imagens interessantes de caracteres (OCR), aviões, peças de máquinas, placas de circuito, mapas, radiografias, órgãos, células, cromossomos, formas de onda, etc., foram analisadas através de características de formato. Sistemas de inspeção visual automática tem despertado muita atenção na indústria pois provê meios econômicos, eficientes e precisos para obtenção de controle de qualidade.

Robôs que “vêem” estão sendo construídos para automatizar a inspeção e manipulação de partes industriais em indústrias manufatureiras (Kruger & Thompson, 1981). Para determinar esta tarefa é necessário que o robô controlado por computador distinga visualmente a forma dos objetos. Tipicamente são empregadas características simples tais como perímetro, área e extensão bidimensional. Os atos dos robôs baseiam-

se no uso de características de forma não preservativas de informações. Estas podem ser definidas como medidas que não permitem uma reconstrução única do objeto original, entretanto fornecem bom indicadores de formato. Muitas tarefas entretanto necessitam que os objetos possam ser reconhecidos de pontos de vistas arbitrários e isso implica na extração de características que sejam invariantes à escala (contrações ou dilatações), translação (vertical e horizontal) e/ou à orientação (rotação em superfície).

Existem muitas teorias de descrição e reconhecimento de Forma, cada uma delas se esforça para explicar algum aspecto do problema. Isto ocorre devido ao fato de podermos conceitualizar a percepção da forma como uma função de alto-nível (*high-level*). Existem muito poucas evidências neurofisiológicas sobre sua natureza e muito do que se conhece baseia-se em hipóteses.

Embora existam muitas abordagens para a Forma, os problemas relacionados à sua descrição e reconhecimento, que pertence tanto a Neurofisiologia e a Psicologia da Percepção, continua longe de solucionar-se. Ainda existem muitas dúvidas como as informações provenientes destas áreas estão relacionadas.

Uma das coisas mais importantes a ser determinada em uma imagem é a forma dos objetos contidos na mesma. Webster define a forma como “a qualidade de um objeto, a qual depende da posição relativa de todos os pontos que compõem seu contorno ou superfície externa”.

Como ressaltado por Lord Kelvin apud Bribiesca & Guzman (1980) uma Ciência emerge quando existe a possibilidade de medir o fenômeno que tal ciência tenta entender, permitindo comparações quantitativas e relações matemática entre elas. Bribiesca & Guzman, (1979) define região como uma porção conectada de um plano limitado por uma borda curva (não há buracos nem auto-intersecção). Uma dada região possui um tamanho, posição e orientação em um plano, isso define uma região plana que é definida unicamente pela curva. Uma forma é o que permanece da região, independente da sua posição, orientação e tamanho no plano. Uma região e sua imagem refletida (*mirror*) não possuem a mesma forma geralmente. Essa definição coincide com a definição psicológica intuitiva de forma. Se uma notação está sendo utilizada para representar a forma da região, ela tem de ser independente da posição, orientação e tamanho de tal região. Uma forma é discreta se o contorno da região é formado por segmentos de uma grade (*grid*) quadrada. O tamanho, orientação e posição da grade afeta a forma discreta resultante.

De acordo com Haber & Wilkinson, (1982), o sistema visual humano é projetado para produzir uma percepção organizada. Uma grande variedade de características espaciais tais como tamanho, forma, distância, posição relativa e textura, são estruturadas pela mente a fim de representar cenas visuais. Estas características espaciais são percebidas como propriedades de coisas, objetos na cena, e não meramente linhas abstratas ou superfícies. Não percebemos linhas ou extensões, percebemos objetos. Todas as partes de cada objeto são percebidas juntas em uma construção global (“*todo*”), não separadas e independentes. Todos os objetos são percebidos como estando relacionados aos outros em relações do tipo *perto*, *longe*, *atrás*, *ao lado*, e assim por diante.

O Reconhecimento da Forma pode ser considerada como resultado de ambos estágios de baixo nível (*low level*) e de alto nível (*high level*) de análise. Segundo Marr, (1976) a extração de uma forma visual pode proceder sua descrição, como consequência, normalmente é mais fácil computar uma descrição menos precisa (*coarse*) da forma antes de ter qualquer idéia sobre qual forma é.

A análise de baixo nível envolvem a agregação de dados imperfeitos de borda na projeção da imagem bi-dimensional.e provavelmente ocorre no córtex visual primário Levine, (1985). Talvez alguns atributos de forma possam ser também computados a partir do conjunto de bordas. Neste nível, pontos de bordas e regiões são importantes mas ainda não estão rotulados, no entanto essas informações servem como entrada para subseqüente processo de alto nível de organização e entendimento.

A coleção de bordas na análise de baixo nível sugere um contorno descrevendo uma área, mas essas regiões podem possuir propriedades similares como cor e textura. Zusne, (1970) sugeriu a seguinte definição: “*contorno é uma interface unidimensional entre a figura e o fundo*”. Esta definição é aceitável restritamente no caso do objeto está sobreposto ao fundo. Geralmente assume-se que um contorno é formado por um conjunto de bordas conectadas umas as outras, mas não necessariamente formando um uma curva fechada.

Zusne, (1970) caracterizou uma hierarquia de definições de generalidade crescente para “forma”, “figura” e “formato”, termos os quais são sinônimos. Na base dessa hierarquia estão os pixels e os elementos de borda, os quais representam o mapa de bordas ou “*primal sketch*”. O próximo nível de complexidade envolve contornos, os quais também circunscrevem regiões. No nível seguinte, coleções de contornos, regiões e

bordas são usadas para modelar objetos no espaço bi-dimensional, e está definido como “domínio da imagem” (*picture domain*). O próximo nível na hierarquia descreve modelos de objetos no espaço tri-dimensional e constitui o “domínio da cena” (*scene domain*). Regiões no domínio da imagem são mapeadas em superfícies para o domínio da cena.

2 – CARACTERÍSTICAS DE FORMA BASEADAS NA TEORIA DOS MOMENTOS

O Reconhecimento da forma dos objetos em uma cena é facilmente determinado pelo sistema visual humano, mesmo quando o objeto está transladado, rotacionado, escalonado, parcialmente obscuro, distorcido ou visto sob diferentes perspectivas. O reconhecimento invariante de formas é importante para seres humanos em diversas tarefas, apesar do reconhecimento *variante* ser importante em algumas tarefas, por exemplo reconhecer “b”, “d”, e “p”. Em Visão de Máquina o reconhecimento invariante da forma é também importante. Em alguns problemas a determinação da posição (transformação de translação) ou orientação (transformação da rotação) de um objeto é primordial. Por exemplo, se um robô necessita pegar um objeto de forma conhecida, a posição e orientação do objeto deve ser conhecida a fim de orientar e posicionar a mão do robô. Em outros casos, a forma do objeto é essencial e não sua posição e orientação. Muitos exemplos dessa situação ocorrem na Inspeção Visual Automática onde algumas propriedades do objeto devem ser medidas e comparadas à outras medidas padrões. Várias medidas exibem propriedades invariantes, entretanto segundo Hall & Park (1986) os Momentos possuem a teoria matemática mais elegante. A aplicação prática dos momentos e seus invariantes são ilustradas em diversas aplicações em sistemas de visão de máquina.

Muitas aplicações envolvem segmentar imagens, localizar regiões de interesse e então determinar a identidade das regiões. Existem duas abordagens principais para descrever a forma de uma figura bidimensional fechada. Uma abordagem é extrair descritores de formato baseados no perímetro da figura (contorno). Quando existe a necessidade de extrair detalhes finos da forma do objeto, os métodos baseados em perímetro são bons, embora sejam muito sensíveis à ruídos. Esse efeito pode ser reduzido utilizando descritores de forma em toda massa da figura, como fazem os momentos.

O método dos Momentos proporcionam uma técnica robusta para a decomposição de uma forma arbitrária em um conjunto de características. O aspecto principal dessa abordagem é que ela baseia-se em transformações lineares diretas sem a necessidade de se determinar parâmetros específicos. Apresentam simplicidade matemática e são muito versáteis, sendo aplicados em vários tipos de imagens e aplicações. A maior limitação dos Momentos é que não são sensíveis à características específicas dos objetos e pode apenas ser aplicado em tarefas de reconhecimento globais. Problemas típicos de todas as características globais são sensibilidade à ruído, os valores mudam drasticamente com oclusões e não são suficientes para distinguir todas as formas.

Momentos são muito úteis para a análise de forma. Momentos de ordem zero e de primeira, segunda e terceira ordens podem fornecer informações úteis em várias aplicações de Visão computacional como localização dos objetos, reconhecimento de formato, orientação e tamanho, além das características invariantes dos momentos quanto a rotação. A forma é frequentemente representada por uma imagem binária e seus momentos podem ser obtidos através de algoritmos rápidos, considerando a forma como um arranjo discreto de pontos, podendo então ser implementado em tempo real (Anderson, 1985).

O problema relacionado com o Reconhecimento de Padrões é decidir como um padrão em branco e preto é “similar” de alguma maneira com um dado protótipo. O padrão pode ser considerado como definido por uma fronteira retangular cobertas de pontos. O que constitui “similaridade” varia muito de caso para caso. Por exemplo pode se considerar que dois padrões são similares somente se eles concordam pixel a pixel, ou pode se definir que dois padrões são topologicamente equivalentes de tal maneira que qualquer curva simples fechada possa ser chamada de similar ao círculo. Muitas definições intermediárias são possíveis. A figura A2.1 a seguir ilustra o conceito de padrões similares utilizados por Alt, (1962).

Figura A2.1 – Exemplo de padrões “similares” – Alt (1962).

As técnicas baseadas em momento descrevem as propriedades globais de uma imagem, isto é, a maneira com que as características de pixels brancos e pretos representam o padrão. Existem outros métodos relacionados com momento onde as características derivam da distribuição estatística dos pontos, como as técnicas de Zonificação (*Zoning*) e N-Tuplas (*N-Tuples*) descritas em Neves (1995).

Existem poucas técnicas que podem ser diretamente comparadas com a abordagem dos Momentos. Descritores de Fourier é uma das técnicas que pode ser comparada aos Momentos para reconhecimento de formas binárias, pois é possível que seja sensível a mais tipos de formato de contorno, entretanto em aplicações práticas é difícil prever qual técnica irá produzir melhores resultados sem realizar experiências empíricas.

Simples transformações nos momentos da imagem podem torná-los invariantes à variáveis que não são de interesse ou são confusas tais como localizações dentro da imagem ou orientações, quando está se estimando o comprimento. A invariância é determinada através da normalização dos momentos de ordens mais alta pelos momentos de ordens mais baixas, os quais indicam posição, orientação, etc. Um exemplo comum de momento invariante é a variância da distribuição de probabilidade, a qual é independente da média ou translação da distribuição. Quando os momentos invariantes são usados para reconhecimento de caracteres, pode-se por exemplo normalizá-los em relação a altura, largura, inclinação, bem como posição e orientação.

3 - APLICAÇÕES

Atualmente diversas áreas tem utilizado procedimentos de Reconhecimento de Padrões e Análise de Imagens, como por exemplo as áreas de processamento de informações escritas e de documentos (OCR). Na área médica temos a identificação e diagnóstico automático de tumores, contagem de células de sangue, sistemas de visualização, etc. Na indústria a inspeção automática é uma das principais aplicações. Na área de segurança temos a identificação de face, da íris e do formato da mão. Em aplicações militares temos a identificação de alvos táticos, interpretação de imagens de satélite, etc. Na robótica diversos sistemas de Visão e Reconhecimento de Padrões têm sido desenvolvidos, seja para aplicações em lugares perigosos aos seres humanos, seja para auxílio a deficientes físicos.

Em geral, funções de várias variáveis podem ser caracterizadas por uma seqüência de números, por exemplo coeficientes de Taylor e coeficientes de Fourier. Momentos representam uma técnica descritiva e intuitiva baseada no estudo da Mecânica dos Corpos. Hu, (1961) desenvolveu a teoria dos Invariantes de Momentos (IM) e sugeriu a sua aplicação para o reconhecimento invariante de padrões visuais, baseado na teoria dos Invariantes Algébricos (Elliott, 1913). O termo “invariante” denota uma quantidade que permanece inalterada sobre certas condições. A teoria dos invariantes algébricos de transformação linear em conexão com uma série de problemas em Geometria Analítica e foi formulada primeiramente no século passado no trabalho de Carley & Sylvester (Gurevich, 1964). Baseado nesse trabalho, Hu apresentou fórmulas para funções invariantes contendo momentos de segunda e terceira ordem, como exemplos da teoria algébrica de invariantes de momentos.

Os invariantes de Momento representam um conjunto completo de descritores de imagem, possuindo portanto uma propriedade de fundamental importância para a teoria de Reconhecimento de Padrões. A teoria dos invariantes de momento tem sido um dos tópicos importantes em muitos livros básicos para Reconhecimento de Padrões como Duda & Hart, (1973), Hall (1979), Ballard & Brown (1982), Gonzales & Wintz (1987), Pratt (1991), Jain (1995).

Desde a sua publicação, os MIs têm sido aplicados no reconhecimento de objetos e na classificação de imagens bidimensionais (silhueta e contorno), 2 ½ D (*range*), 3D e em níveis de cinza (intensidade) (Prokop & Reeves, 1992), como por exemplo no reconhecimento de:

- Caracteres impressos (Giuliano et al. 1961¹; Alt, 1962; Lambert, 1969; Cash & Hatamian, 1987; Wong *et al*, 1995, manuscritos (Casey, 1970)
- Classificação de histogramas bidimensionais de amplitude e freqüência de diferentes modulações em sinais de HF (Smith, 1967)
- Identificação de barcos (Smith & Wright, 1971)
- Imagens de Raio-X (Hall *et al*, 1975)
- Comparação de cenas (Wong & Hall, 1978; Maitra, 1979)
- Detecção de alvos (prédios e pontes) (Gilmore & Boyd, 1981)

¹ Trabalho realizado em paralelo ao de Hu (1961).

- Identificação de aviões (Dudani *et al.*, 1977; Reeves & Taylor, 1989, Belkasin *et al.*, 1991)
- Classificação de espécies de peixes (Strachan *et al.* 1990)
- Determinação da orientação de estruturas tridimensionais em imagens médicas (Faber & Stokely, 1988).
- Inspeção automática de material impresso (Simpson *et al.*, 1991)
- Reconhecimento de mapas e ferramentas (Chen, 1993)
- Peças Mecânicas (Gonzaga & Costa, 1996)
- Fotografias sem nitidez (*blurred images*) (Flusser *et al.*, 1996)
- Detecção de câncer de mama (Chang *et al.*, 1998)
- Armazenamento e recuperação de logotipos em base de dados (Jain & Vailaya, 1998)

Giuliano *et al.*, (1961) realizaram uma das primeiras experiências de reconhecimento de caracteres com Momentos, baseando-se no *Método dos Eixos Principais*, entretanto houve pouca pesquisa subsequente baseada em seu método. O método dos eixos principais baseia-se na observação de que momentos podem ser calculados em relação aos eixos principais da distribuição e será portanto invariante a orientação da distribuição, entretanto o método falha para figuras rotacionalmente simétricas, onde o conjunto de eixos principais não é único.

Alt, (1962) determinou o reconhecimento de caracteres impressos invariantes à localização, tamanho, compressão e expansão, inclinação (tipo itálico), pequenas mudanças de proporções, curvatura, etc. e à uma certa quantidade de ruído, através da transformação dos momentos. Através das características invariantes cria-se um conceito de similaridade definida por uma lista de transformações as quais não afetam a similaridade entre os padrões.

Casey (1970) usou os momentos de segunda ordem para especificar as transformações das coordenadas das formas de onda. Caracteres em letra de forma são transformados linearmente a fim de melhorar a aparência para reconhecimento (normalização). Através do uso de momentos, o padrão original foi mapeado para um domínio livre de variações e as variações lineares dos caracteres manuscritos foram removidas. Isso indica que através da normalização realizada pelos momentos após digitalização pode reduzir a taxa de erros.

Em Nagy & Tuong (1970) são aplicadas técnicas de normalização de padrões baseadas na projeção geométrica, com o objetivo de reduzir cada padrão a uma forma canônica e aumentar a similaridade entre padrões de uma mesma classe.

Em Naylor, (1971) é descrita uma técnica para normalização da inclinação de caracteres manuscritos, através de transformadas.

Em Smith & Wright (1971) as imagens dos barcos foram geradas, especificando-se o contorno dos barcos e sua densidade média. Uma grade de células quadradas é deslocada sobre o barco e para as células com os centros dentro ou no contorno eram atribuídos valores 0 ou 1, com a probabilidade correspondente à densidade média. O contorno do navio é considerado estar inscrito em um retângulo e ser simétrico em relação ao seu eixo principal.

Wong & Hall (1978) utilizaram um conjunto de invariantes de momento para realizar *scene matching* de imagens aéreas provenientes de radar. Devido às propriedades invariantes, é fácil se determinar um *threshold* e um *match* eficiente. Sub-regiões quadrulares da imagem de radar foram comparadas com sub-regiões da imagem de radar através de uma correlação baseada no log dos valores dos momentos (para reduzir a faixa dinâmica de valores). Os momentos invariantes mostraram-se muito úteis para realizar um *matching* entre imagens entretanto o tamanho entre as imagens ópticas e do radar foram consideradas de mesma escala e orientação.

Sadjadi & Hall (1978) investigaram a eficiência dos invariantes de momento para análise de cenas. Através de uma experiência simples, eles demonstraram a consistência da aplicação de momentos em imagens em níveis de cinza, sujeitas a várias transformações de tamanho e rotação. Os sete invariantes de Hu mostraram-se similares em todas as imagens transformadas.

Teague (1980) e Casasent & Psaltis (1980) propuseram esquemas que aplicam técnicas de processamento ópticas ou híbridas para gerar os invariantes de momento automaticamente para fins de reconhecimento de padrões.

Gilmore & Boyd (1981) utilizaram os sete invariantes de Hu para identificar construções bem definidas e pontes alvos. Nesse trabalho, a orientação e a variação de intensidade (range) do sensor de imagem são conhecidos de maneira que a forma

esperada e o tamanho do alvo podiam ser calculados com base no modelo do alvo. Primeiro a cena foi segmentada e delimitada (*thresholded*) em diferentes regiões de silhuetas. Uma etapa de pré-processamento foi realizada para desqualificar regiões que diferem muito do alvo esperado. A classificação foi baseada na diferença ponderada entre os momentos das regiões e os momentos esperados para o alvo, obtendo uma classificação correta para os alvos.

Abu-Mostafa & Psaltis (1984) avaliaram os invariantes de momento como um espaço de características para Reconhecimento de Padrões em termos de poder discriminante e tolerância à ruídos.

Cash & Hatamian (1987) utilizaram momentos centrais normalizados em relação à escala (ℓ_{00} , ℓ_{20} , ℓ_{02} , ℓ_{11} , ℓ_{21} , ℓ_{12} , ℓ_{30} , ℓ_{03}) para extrair características de caracteres em um sistema aplicados à seis fontes impressas. Os resultados indicaram que os momentos forneceram boa discriminação entre caracteres e foram obtidas taxas de reconhecimento próximas a 100%. Utilizando-se apenas os momentos centrais, a informação de altura e largura do objeto é perdida (os momentos ℓ_{10} e ℓ_{01} são iguais a zero), assim os momentos m_{10} e m_{01} foram incluídos como características, resultando em um aumento na taxa de reconhecimento para todas as fontes. Foram testados três métodos diferentes de classificação: Distância Euclidiana, Correlação Cruzada e Distância de Mahalanobis.

Strachan et al. (1990) compararam três métodos diferentes para classificação de espécies de peixes baseados em invariantes de momento de Hu, otimização de comparação de matrizes e um outro descritor de formato que baseia-se na distância entre as bordas do peixe tomadas em 10 posições fixas em função do comprimento do eixo principal. Os resultados obtidos para os métodos foram respectivamente 73%, 63% e 90%.

Kothanzad & Lu, (1990) utilizaram como classificador um *perceptron* de duas camadas treinados com *back-propagation*, utilizando Invariantes de Momentos e Momentos de Zernike em imagens de caracteres manuscritos e verificaram que sua performance foi superior a três outros classificadores tradicionais: vizinho mais próximo (*nearest neighbor*), Bayes e distância mínima média, especialmente quando ruído estava presente. Segundo eles, os Momentos de Zernike possuem um grande poder de separabilidade entre as classes e são mais poderosos que os IMs de Hu.

Em Simpson et al. (1991) os invariantes de momento são utilizados para a detecção de defeitos em inspeção automática de material gráfico impresso e também foram investigados a sensibilidade dos invariantes de momento em relação a localização dos defeitos, área, forma e densidade. O potencial de utilizar os momentos para separar materiais defeituosos de materiais não-defeituosos foi demonstrada, entretanto a digitalização e ruídos alteram a performance. Os invariantes de momento ainda apresentam atributos mais atrativos em relação às técnicas convencionais para inspeção de material impresso, tais como subtração de imagens.

Zhou *et al.*, (1991) apresentaram um modelo de atenção seletiva para reduzir o número de entradas da rede e reconhecer imagens contendo múltiplos objetos distorcidos, em qualquer posição e escala. A invariância à posição e escala é obtida é conseguida através da representação da imagem por atributos invariantes e normalização da posição e orientação padrões.

A maioria dos trabalhos descritos acima estão limitados a aplicação dos invariantes de momento de baixa ordem (invariantes de Hu). Fórmulas explícitas para funções invariantes para ordens superiores praticamente não haviam sido apresentadas. Li (1992) apresentou uma nova abordagem para calcular momentos invariantes diferente da abordagem de Hu e pode ser usada para derivar funções invariantes de uma ordem arbitrária por uma combinação de componentes no espectro de Fourier-Mellin da função de distribuição da irradiância da imagem. São apresentadas cinquenta e duas funções (de segunda a nona ordem), invariantes à translação, escala e rotação e que podem ser utilizadas como funções de bibliotecas em programas de computador para identificação automática de imagens.

Sugisaka & Teshnehlal, (1993) apresentaram um sistema para reconhecimento dinâmico de padrões, combinando os invariantes de momento e uma Rede Neural de três camadas (com 50 unidades na camada intermediária). Resultados obtidos para o reconhecimento de quatro tipos diferentes de padrões (forma), em três tamanhos e diferentes posições são apresentados.

Gomes & Carvalho Filho, (1994) utilizaram momentos e rede neural de múltiplas camadas para descrever assinaturas armazenadas em um banco de imagens.

Wong *et al.*, (1995) apresentaram uma nova forma de derivar automaticamente os IM e diversos novos IMs foram derivados. Foram discutidos diversos fatores relativos à discriminação dos atributos e foram obtidas taxas de reconhecimento de até 98.8% para uma aplicação de OCR.

A melhor maneira de se otimizar o grande tempo gasto no cálculo dos momentos é associá-los a um *hardware* especializado. A nível de *software* essa otimização vem sendo realizada através do desenvolvimento de algoritmos paralelizados que permitem que várias regiões da imagem sejam analisadas simultaneamente.

Computações eficientes de momentos utilizando processadores ópticos digitais (Casasent & Psaltis, 1980, Teague, 1980, Casasent et al., 1982) ou filtros (Budrikis & Hatamian, 1984) representam uma vantagem importante em aplicações para reconhecimento de forma em tempo real.

Casasent & Psalts (1980) descreveram um processador híbrido óptico-digital que calcula todos os momentos em paralelo. A luz de um *laser* passa através de uma transparência da imagem $f(x,y)$ e então através de uma máscara $g(x,y)$ e então através de uma lente de transformada de Fourier e o padrão final é coletado em um fotodetector. A seleção de uma máscara causa na saída um momento específico, portanto diferentes momentos necessitam de diferentes máscaras.

Em Revees, (1982) pode ser vista uma computação paralelizada de propósito geral para o cálculo dos momentos. Vários outros métodos que reduzem a complexidade computacional do cálculo dos momentos têm sido propostos como em Zakaria et al. (1987)² e Li (1993)³.

Hatamian (1986) propôs um algoritmo e um único *chip* para implementação VLSI para gerar momentos na taxa de vídeo. Segundo Hatamian 16 momentos podem ser computados a uma taxa de 30 *frames* /s, em uma imagem de 512x512x8 bits. O algoritmo dos momentos é baseado no uso de uma função geradora de momentos discreta

²Esse método denominado “Método Delta” calcula os momentos de imagens contíguas utilizando a contribuição de cada linha ao invés da contribuição individual de cada pixel. O Método Delta necessita da coordenada do primeiro pixel do objeto e o comprimento da sequência de pixels do objeto (sequência de 1s), dessa maneira o custo computacional pode ser reduzido ao perímetro do objeto. Nesse trabalho, os autores estimaram um tempo inferior a 0.6 s para calcular dos invariantes de Hu para uma matriz 80 x 640 .

unidimensional como um filtro digital. A transformada Z analisa as respostas desses filtros e a implementação consiste em um arranjo de filtros bidimensionais.

Dai et al (1992) propuseram um interessante algoritmo para calcular os momentos considerando as contribuições apenas dos pontos de borda. O método denominado Método Integral é derivado do método de Li (1993) e utiliza um algoritmo recursivo a fim de evitar multiplicações e pode ser utilizado para calcular momentos a partir de *run-length codes*⁴ ou *chain codes*.

Chen, (1993) propôs um método para melhorar a performance no cálculo dos momentos baseando-se também apenas em pontos de borda, através do cálculo da integral de linha ao longo da curva⁵.

Yazdi et al, (1993) apresentou uma arquitetura VLSI de uma rede neural de múltiplas camadas como classificador e como características os invariantes de momento calculados a partir dos contornos dos objetos.

Em Cheng et al., (1998) também é mostrado o desenvolvimento de uma arquitetura VLSI para o cálculo de momentos centrais e momentos regulares. Características de momento foram extraídas de imagens de biópsias de câncer de mama e classificadas por uma rede neural, determinando uma taxa de classificação de 100%.

Conforme procurou-se mostrar nos trabalhos revistos acima, o uso de momentos tem se tornado um método muito útil e de grande aplicabilidade para o Processamento de Imagens, Reconhecimento de Padrões e Visão por Computador.

³ Segundo Li, esse método é duas vezes mais rápido do que Zakaria et al (1987).

⁴ Qualquer região ou forma pode ser vista como uma seqüência (*string*) alternadas de 0s e 1s. O código da *string* (*run-length code*) consiste em endereçar cada *string* de 1s, seguida de seu comprimento (número de 1s). Esse método possui a vantagem de representar a imagem em uma única varredura, independente da complexidade da imagem, mas sua principal desvantagem é que não fornece os pontos das bordas ordenados ao longo do seu contorno como no *chain code* (código da cadeia) e isso torna difícil a segmentação se várias regiões diferentes estão presentes na imagem.

⁵ Com esse método, Chen estimou que o tempo gasto no cálculo dos sete invariantes de momento para uma imagem de 250x250 pixels necessita de 0.6 s em comparação aos 15 s gastos no método convencional.

4 – DEFINIÇÕES

Momentos descrevem quantidades numéricas à alguma distância de um ponto ou eixo de referência. Momentos são comumente utilizados em Estatística para caracterizar a distribuição de variáveis randômicas (Papoulis, 1965) e similarmente em Mecânica para caracterizar os corpos através da distribuição espacial de suas massas. O uso de momentos para análise de imagem é direto se consideramos uma imagem binária ou em nível de cinza segmentada como uma função de distribuição de densidade bidimensional, desta maneira os momentos podem ser usados para caracterizar segmentos de imagem e extrair propriedades que possuem analogia com a Estatística e a Mecânica. Esse método é simples de calcular e requer apenas um passo em toda a imagem. Normalmente os descritores de formato estão restritos aos momentos de ordem mais baixas onde os erros de amostragem em objetos digitalizados é maior.

As imagens são normalmente representadas por funções bidimensionais de intensidade de luz $f(x,y)$, onde x e y são as coordenadas espaciais cartesianas e o valor da função f em qualquer ponto (x,y) é proporcional ao brilho da imagem naquele ponto (luminância).

O teorema da *Unicidade* estabelece que existem momentos de todas as ordens se $f(x,y)$ for contínua e possuir valores não nulos em apenas uma região do plano cartesiano (x,y) (Hu, 1962). O conjunto de momentos $\{m_{p,q}\}$ é determinado de uma maneira *única* por $f(x,y)$ e $f(x,y)$ é determinada unicamente por $\{m_{p,q}\}$.

Para representar toda a informação em uma região de uma imagem é necessário um número infinito de valores de momentos. Na prática isso é impossível, assim um dos objetivos no projeto de um sistema de reconhecimento que utilize essa técnica é a definição da ordem n dos momentos. Essa ordem deve fornecer informação suficiente para representar a imagem de uma maneira única, respeitando-se os critérios de velocidade de reconhecimento e o poder discriminante dos padrões a serem classificados.

A definição de momento⁶ (ou algébricos) geométricos regulares tem a forma de uma projeção da função $f(x,y)$, que representa a imagem, em uma função monomial $x^p y^q$. O momento $(p+q)$ é definido pela equação (1) a seguir:

⁶ Um momento é uma soma ponderada de um padrão $f(x)$ sobre todo o campo de entrada, com os pesos iguais à algum polinômio em x (Wood, 1996)

$$m_{p,q} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^p y^q \cdot f(x,y) \cdot dx dy \quad (1) \quad (p, q = 0, 1, 2, \dots)$$

A definição acima tem a forma de uma projeção da função $f(x,y)$ dentro do monômio $x^p y^q$, entretanto a base embora completa não é ortogonal. O cálculo de m_{pq} consiste em se multiplicar a função $f(x,y)$ por um monômio $x^p y^q$ e integrar o resultado. Os monômios até ordem 3 ($x^0 y^0$, $x^0 y^1$, $x^0 y^2$, $x^1 y^0$, $x^1 y^1$, $x^1 y^2$, $x^2 y^0$, $x^2 y^1$, $x^2 y^2$, $x^3 y^0$, $x^3 y^1$, $x^3 y^2$) podem ser vistos na figura A2.2 Os momentos de ordem $(p+q)$ também podem ser interpretados como a resposta de um sistema de formação de imagem com uma função de transferência $x^p y^q$ e entrada $f(x,y)$. Os momentos de baixa ordem possuem relações intuitivas com os objetos, por exemplo m_{00} refere-se à massa, m_{10} e m_{01} referem-se ao centro de massa e m_{11} , m_{20} e m_{02} referem-se às direções principais.

As generalidades do método dos momentos e o uso dos invariantes de momento são descritos em Deutsch, (1955). Ele propôs uma teoria do processo de percepção humana que baseia-se na hipótese de que os seres humanos medem algum tipo de informação baseada em momento nas imagens. Além disso, ele propôs que a informação contida na relação de um ponto com seus vizinhos demonstra invariância de tamanho, localização, orientação, reflexão e simetria, identificadas também nos experimentos de percepção humana.

Deutsch, (1955) descreveu quatro propriedades invariantes da percepção humana:

- 1) - Formas devem ser reconhecidas independentes de suas localizações no campo visual. Não é necessário fixar os olhos no centro da figura para reconhecê-la, nem mover os olhos ao longo de seu contorno, assim o reconhecimento de forma é invariante à translação.
- 2) - O reconhecimento pode ser efetuado independentemente do ângulo de inclinação da figura no campo visual. O ângulo refere-se à orientação angular da figura em duas dimensões embora tenha sido observada alguma invariância ao ângulo de profundidade. O reconhecimento de uma figura é invariante à rotação em duas dimensões, embora as vezes seja necessário se reconhecer o ângulo de rotação.
- 3) O tamanho de uma figura não interfere no reconhecimento de sua forma, assim o reconhecimento visual exibe invariância à escala.

- 4) Imagens refletidas (espelho) parecem diferente. Seres humanos e animais tendem a confundí-las.

Deutsch (1955) destacou ainda que certos organismos primitivos acham mais difícil distinguir entre círculos e quadrados do que entre retângulos e quadrados. Ele também argumentou que as propriedades de invariância estão presentes em toda parte do córtex visual. A simplicidade e potencial da abordagem de Deutsch representou uma das principais contribuições da psicologia no entendimento de sistemas de reconhecimento.

Figura A2.2 – Monômios de até terceira ordem (x^0y^0 , x^0y^1 , x^0y^2 , x^1y^0 , x^1y^1 , x^2y^0 , x^2y^1 , x^3y^0 , x^1y^2) que multiplicam a função $f(x,y)$ no cálculo dos momentos (Sadjadi & Hall, 1978).

O momento bidimensional para uma imagem $f(x,y)$ discreta de tamanho $(N \times M)$ é:

$$m_{p,q} = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} x^p y^q \cdot f(x,y) \quad (2) \quad (p, q = 0, 1, 2, \dots)$$

Esses momentos são chamados de momentos convencionais *ou raw moments*.

A fim de se trabalhar com imagens de qualquer tamanho, pode-se normalizar as coordenadas x e y entre $-1 \leq x, y \leq 1$.

Para normalizar os momentos e torná-los invariantes em relação à translação basta calculá-los em relação ao centro de massa da figura, por isso esses momentos são chamados de *Momentos Centrais*. Para normaliza-los em relação à escala é necessário normalizá-los em relação à área da figura.

Muitas aplicações industriais as imagens podem ser representadas em preto e branco e apenas informações de tamanho, contorno, semelhança e contiguidade são importantes, não a cor ou a sombra dos objetos. Nesse caso a função intensidade $f(x,y)$ consiste apenas de 0s e 1s.

As técnicas de normalização de tamanho, orientação, posição, inclinação ou espessura das linhas visam facilitar o manuseio dos dados e a extração de características, melhorar a taxa de reconhecimento em relação a invariância do sistema em relação a estes atributos.

Outros parâmetros como o ângulo do eixo principal da figura, determinado pelos momentos, podem ser utilizados para normalizar em escala, orientação e posição os padrões contidos na imagem, para efeitos de comparações com imagens e parâmetros dos objetos previamente armazenados em bancos de dados.

Em Weiss (1993) pode ser vista uma discussão sobre o conceito geral de invariância no reconhecimento de objetos e apresenta algumas técnicas clássicas da literatura em invariância projetiva. Wood (1996) apresenta uma revisão sobre vários métodos envolvidos no reconhecimento de padrões invariantes. Em Costa (1996) é mostrada uma revisão de reconhecimento de padrões invariante utilizando Redes Neurais Artificiais (RNA).

5 – INVARIÂNCIA À TRANSLAÇÃO E ESCALA

Esses momentos não são invariantes à quaisquer distorções. Pode-se normalizar a equação (2) para tornar os momentos invariantes à mudanças de translação e escala. Para normalizar os momentos em relação à translação basta calculá-los em relação ao centro de massa da figura, por isso esses momentos são chamados de *Momentos Centrais*. Para torná-los invariantes à escala é necessário normalizá-los em relação à área da figura, conforme mostra a equação (3).

$$\mu_{pq} = \frac{1}{m_{0,0}^{1+(p+q)/2}} \sum_{y=0}^{M-1} \sum_{x=0}^{N-1} (x - \bar{x})^p (y - \bar{y})^q \cdot f(x, y) \quad (p, q = 0, 1, 2, \dots) \quad (3)$$

onde $(p+q)$ é a ordem do momento e $m_{0,0}$ (momento de ordem 0) representa a área da região da imagem, como pode ser visto equação (4).

$$m_{0,0} = \sum_{y=0}^{M-1} \sum_{x=0}^{N-1} f(x, y) \quad (4)$$

As equações (5) e (6) representam são as coordenadas de centro de gravidade do objeto na imagem, onde $m_{1,0}$ e $m_{0,1}$ representam os momentos de primeira ordem:

$$\bar{x} = \frac{m_{1,0}}{m_{0,0}} \quad (5)$$

$$\bar{y} = \frac{m_{0,1}}{m_{0,0}} \quad (6)$$

O conjunto completo dos momentos de ordem n são todos os momentos $m_{p,q}$, com $(p+q) \leq n$ e $\frac{1}{2} \cdot (n+1) \cdot (n+2)$ elementos.

Os momentos Centrais são definidos em termos dos momentos ordinários, conforme pode ser visto nas equações (7):

$$\mu_{0,0} = m_{0,0} = \mu \quad (7.a)$$

$$\mu_{1,0} = \mu_{0,1} = 0 \quad (7.b)$$

$$\mu_{2,0} = m_{2,0} - \mu \bar{x}^2 \quad (7.c)$$

$$\mu_{1,1} = m_{1,1} - \mu \bar{x} \bar{y} \quad (7.d)$$

$$\mu_{0,2} = m_{0,2} - \mu \bar{y}^2 \quad (7.e)$$

$$\mu_{3,0} = m_{3,0} - 3m_{2,0}\bar{x} + 2\mu\bar{x}^3 \quad (7.f)$$

$$\mu_{2,1} = m_{2,1} - m_{2,0}\bar{y} - 2m_{1,1}\bar{x} + 2\mu\bar{x}^2\bar{y} \quad (7.g)$$

$$\mu_{1,2} = m_{1,2} - m_{0,2}\bar{x} - 2m_{1,1}\bar{y} + 2\mu\bar{x}\bar{y}^2 \quad (7.h)$$

$$\mu_{0,3} = m_{0,3} - 3m_{0,2}\bar{y} + 2\mu\bar{y}^3 \quad (7.i)$$

ou genericamente como dado pela expressão (8):

$$\mu_{pq} = \frac{m_{pq}}{m_{00}^{\frac{(p+q)}{2}+1}} \quad (8)$$

Vários atributos podem ser determinados a partir destes momentos centrais conforme será mostrado nas próximas seções. As informações sobre o centro de massa, área e orientação do eixo principal mais próximo ao eixo x (θ) podem servir tanto para um manipulador como um braço mecânico que irá manipular os objetos, ou para outro estágio mais elevado, onde as relações entre os diferentes objetos contidos na imagem proporcionam a interpretação da cena.

Um conjunto finito de momentos pode ser armazenada como uma matriz, como por exemplo mostra a matriz M abaixo, composta pelos momento não-centrais com $p, q \leq 3$.

$$M = \begin{bmatrix} m_{00} & m_{01} & m_{02} & m_{03} \\ m_{10} & m_{11} & m_{12} & m_{13} \\ m_{20} & m_{21} & m_{22} & m_{23} \\ m_{30} & m_{31} & m_{32} & m_{33} \end{bmatrix} \quad (9)$$

Se uma ordem $n=p+1$ for seleccionada e todos os momentos com n menor ou igual a um valor fixo forem seleccionados, então é produzida uma matriz triangular superior. Matrizes similares podem ser produzidas para os momentos centrais e momentos normalizados.

Uma questão importante na implementação computacional dos momentos invariantes é a faixa dos momentos uma vez que isso determina os requisitos de armazenamento. É suficiente para implementação computacional usar limites superiores para o valor máximo. Para uma região retangular as bordas da região pode ser facilmente calculadas através da seguinte relação:

$$m_{pq} = \int_a^b \int_c^d x^p y^q \cdot f(x, y) dx dy \leq f \max \int_a^b \int_c^d x^p y^q dx dy \quad (10)$$

o que fornece

$$m_{pq} \leq f \max \left[\frac{b^{p+1} - a^{p+1}}{p+1} \right] \left[\frac{d^{q+1} - c^{q+1}}{q+1} \right] \quad (11)$$

Momentos centrais apresentam limites superiores similares. Limites inferiores podem ser calculados de maneira análoga. Os limites computados para os momentos podem ser usados para escalonar as operações do computador, seleccionar comprimento dos registos e normalizar valores das características computadas (Sadjadi & Hall, 1978).

6 - INVARIÂNCIA À ROTAÇÃO

Muitas aplicações necessitam que os descritores de forma sejam invariantes mas em outras aplicações a posição rotacional de algum objeto pode ser uma informação importante. O problema de se reconhecer objetos invariantes à rotação consiste no fato de que algumas formas apresentam simetrias similares em diversas posições rotacionais. Hu (1962) propôs dois métodos diferentes para se calcular momentos invariantes à rotação: Invariantes de Momento e Eixo Principal. Os momentos de segunda ordem definem eixos principais da mesma maneira que os momentos de inércia na Mecânica.

6.1 – Invariantes de Momento de Hu

O conjunto dos invariantes de momento absolutos consistem de combinações não lineares aplicados aos momentos gerando funções invariantes a transformação de rotação. Através dos momentos de segunda e terceira ordem, Hu definiu sete funções denominadas Invariantes de Momento de Hu (IM) (Hu, 1962), que são invariantes à translação, rotação e escala. Essas funções são mostradas pela equação 12 a seguir.

$$\phi_1 = \mu_{20} + \mu_{02} \quad (12-a)$$

$$\phi_2 = (\mu_{20} - \mu_{02})^2 + 4\mu_{11}^2 \quad (12-b)$$

$$\phi_3 = (\mu_{30} - 3\mu_{12})^2 + (3\mu_{21} - \mu_{03})^2 \quad (12-c)$$

$$\phi_4 = (\mu_{30} + \mu_{12})^2 + (\mu_{21} - \mu_{03})^2 \quad (12-d)$$

$$\phi_5 = (\mu_{30} - 3\mu_{12})(\mu_{30} + \mu_{12}) \cdot [(\mu_{30} + \mu_{12})^2 - 3(\mu_{21} + \mu_{03})^2] + \\ + (3\mu_{21} - \mu_{03})(\mu_{21} + \mu_{03}) \cdot [3(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2] \quad (12-e)$$

$$\phi_6 = (\mu_{20} - \mu_{02})[(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2] + 4\mu_{11}(\mu_{30} + \mu_{12})(\mu_{03} + \mu_{21}) \quad (12-f)$$

$$\phi_7 = (3\mu_{21} - \mu_{03})(\mu_{30} + \mu_{12})[(\mu_{30} + \mu_{12})^2 - 3(\mu_{21} + \mu_{03})^2] - \\ - (\mu_{30} - 3\mu_{12})(\mu_{21} + \mu_{03})[3(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2] \quad (12-g)$$

Os seis primeiros momentos ϕ_1 a ϕ_6 são invariantes à rotação e a reflexão (*absolute orthogonal invariants*), entretanto ϕ_7 é sensível à reflexão (*skew orthogonal invariant*). Seu valor muda de sinal para a imagem refletida de um objeto, mas sua magnitude permanece inalterada. Portanto, caso deseje-se invariância à reflexão apenas

os seis primeiros invariantes de Hu devem ser usados, caso contrário, seu sinal deve ser levado em consideração. O sétimo invariante é útil para distinguir imagens refletidas (*mirror*).

Um padrão desconhecido pode ser classificado por exemplo calculando-se os dois primeiros invariantes de momento e procurando a mínima distância Euclidiana entre eles. Se a distância mínima não está dentro de um *threshold* específico, o padrão desconhecido pode ser considerado uma nova classe, dada uma identidade e adicionado a padrões conhecidos. Uma experiência semelhante foi realizada por Hu (1962) utilizando um conjunto de 26 letras maiúsculas como padrões de entrada. Todos os pontos representando cada um dos caracteres apresentaram-se distintos quando plotados em um espaço bidimensional, entretanto Hu observou que alguns caracteres muito diferentes em forma estavam próximos no espaço de características e que pequenas variações nas imagens de entrada do mesmo carácter resultaram em valores variáveis que se sobrepueram às classes mais próximas. Hu concluiu que o aumento da resolução da imagem e que um espaço de características maior deveria melhorar a distinção entre as classes.

Sabe-se que os momentos que são utilizados para construir os invariantes de momento são em geral definidos na forma contínua mas na implementação prática, são computados de maneira discreta. Observou-se que o cálculo dos momentos de ordens superiores na forma discreta convencional resulta em alguns erros indesejáveis. Lin & Wang (1994) propuseram uma fórmula discreta para computar os momentos mais precisamente. A aplicação da fórmula correta para calcular os invariantes de momento de Hu mostrou que quando utilizados juntos, não há efeito nos últimos seis invariantes devido a cancelamentos dos erros, entretanto ocorre um erro com primeiro e mais importante momento de Hu. As equações (13) abaixo mostra a relação entre os momentos m_{pq} e os momentos corrigidos propostos por Lin & Wang ($\hat{m}_{pq}$). A correção para outros momentos de ordens mais altas também são mostrados nesse trabalho.

$$\begin{aligned}
m_{00} &= \hat{m}_{00} \\
m_{01} &= \hat{m}_{01} \\
m_{10} &= \hat{m}_{10} \\
m_{20} &= \hat{m}_{20} + \frac{1}{12} \hat{m}_{00} \\
m_{02} &= \hat{m}_{02} + \frac{1}{12} \hat{m}_{00} \\
m_{30} &= \hat{m}_{30} + \frac{1}{4} \hat{m}_{10} \\
m_{03} &= \hat{m}_{03} + \frac{1}{4} \hat{m}_{01}
\end{aligned} \quad (13)$$

As variáveis de Hu tendem a ser muito pequenas e possui um enorme range dinâmico. Master, (1994) sugere um conjunto de transformações que atenuam esses problemas e podem ser vistos na equação 14 mostrada a seguir:

$$\begin{aligned}
nh1 &= 10 \log(10\phi_1 - 1.2) \\
nh2 &= 3 \log\left(\frac{\phi_2 + 0.0001}{0.0101}\right) \\
nh3 &= 8 \log\left(\frac{\phi_3 + 0.0005}{0.0025}\right) \\
nh4 &= 2 \log\left(\frac{\phi_4 + 0.0000001}{0.0000101}\right) \\
nh5 &= 1000\sqrt[3]{\phi_5} \\
nh6 &= 200\sqrt[3]{\phi_6} \\
nh7 &= 500\sqrt[3]{\phi_7}
\end{aligned} \quad (14)$$

Uma outra normalização proposta por Khotanzad & Lu, (1990) foi o uso do logaritmo dos módulos dessas funções como atributos dessas funções.

Enquanto os IM possuem algumas propriedades que boas características de formato devem ter, nem sempre são usados todos eles na solução de um determinado problema em particular. Em alguns casos, poucos IM de baixa ordem podem refletir em características discriminantes de formato suficientes (Castleman, 1979).

O método dos eixos principais descrito por Hu fornece as relações rotacionais entre os momentos de mesma ordem. Usando essas expressões o conjunto de momentos de um objeto pode ser rotacionado sobre seu centro de gravidade. A vantagem dos momentos sobre outras técnicas é que esses descritores de formato tem uma implementação direta e também carregam uma interpretação “física” da forma.

Sadjadi & Hall (1978) avaliaram o uso de invariantes de momentos em uma imagem aérea. A imagem foi rotacionada 2° e 45° , seu tamanho foi reduzido pela metade e também foi feita sua reflexão. Os invariantes de momento de Hu extraídos são mostrados na tabela A2.1.

Tabela A2.1 – Momentos Invariantes computados por Wong & Hall, (1978)

Invariante (Log)	Original	Metade do tamanho	Reflexão	Rotação 2°	Rotação 45°
ϕ_1	6.249	6.226	6.919	6.253	6.318
ϕ_2	17.18	16.954	19.955	17.27	16.803
ϕ_3	22.655	23.531	26.689	22.836	19.724
ϕ_4	22.919	24.236	26.901	23.13	20.437
ϕ_5	45.749	48.349	53.724	46.136	40.525
ϕ_6	31.83	32.916	37.134	32.068	29.315
ϕ_7	45.589	48.343	53.590	46.017	40.470

6.2 – Momentos Invariantes de Hu (Método dos Eixos Principais)

Uma outra metodologia para a normalização em relação à rotação consiste em rotacionar o objeto para uma posição padrão, método que foi definido por Hu, (1962) como Momentos Invariantes (MI). Nesse método os momentos são calculados em referência a um par de eixos principais determinados unicamente. A maneira mais fácil de se definir uma posição padrão para a figura é achar o seu eixo principal, ou seja, a direção onde ocorre o maior espalhamento, e rotacioná-lo até que seja horizontal. Existem duas orientações que satisfazem esse requisito, uma deslocada 180° da outra. Essa ambigüidade pode ser resolvida fazendo com que a parte mais pesada da figura fique no lado esquerdo. Na linguagem estatística dizemos que o objeto está inclinado à direita.

Como exemplo, a figura A2.3 abaixo, extraída de Master, (1994) mostra a posição padrão.

Figura A2.3 – Posição padrão para Invariância à Rotação – Master (1994).

Essas condições podem ser estabelecidas em termos de seu efeito nos momentos centrais. Para satisfazer as condições acima, os momentos da figura deve possuir as seguintes propriedades:

- $\mu_{11}=0$. Essa condição se verifica quando os eixos principais da figura estão alinhados com os eixos das coordenadas vertical e horizontal.
- $\mu_{20} > \mu_{02}$. Essa condição estabelece que o eixo principal da figura é horizontal.
- $\mu_{30} > 0$. Essa condição assegura que a maior massa do objeto encontra-se do lado esquerdo, com sua cauda inclinada à direita.

Essas condições estabelecidas acima nem sempre podem ser obtidas ou também podem não ser únicas, embora na maioria das vezes resultem em uma única orientação. A primeira condição pode ser satisfeita no mínimo em quatro posições rotacionais. Isso ocorre quando os dois eixos principais da figura são horizontais e verticais. Algumas simetrias podem resultar em mais que quatro posições. A Segunda condição estabelece que o comprimento do eixo horizontal (rotacionado) é maior que o eixo vertical. Se os eixos principais forem exatamente do mesmo tamanho, como ocorre por exemplo nos quadrados e círculos, esse método não é apropriado pois não é possível achar a posição padrão. A terceira condição é incluída para eliminar a ambigüidade final pois uma figura cujo eixo principal foi rotacionado para a direção horizontal, se rotacionada 180° , continuaria sendo horizontal. Neste caso, deseja-se que a figura não tenha simetria esquerda-direita pois neste caso o terceiro momento horizontal será zero.

O primeiro passo para a padronização da posição da figura consiste em determinar a direção de seu eixo principal, isto é, a direção de maior espalhamento da figura. Para algumas formas como o círculo, a orientação não é única. Para se definir uma única orientação é necessário que o objeto seja alongado, dessa forma a orientação de seu eixo de alongação pode ser usado para definir sua orientação. Usualmente o eixo de menor segundo momento, que em 2-D é equivalente ao eixo de menor inércia, é usado como eixo de alongação. O eixo de segundo momento de um objeto é a linha na qual a soma das distâncias ao quadrado entre pontos do objeto e a linha é mínima. Conforme demonstrado em Jain et al, (1995) o eixo de orientação é obtido minimizando-se a expressão mostrada na equação (15):

$$X^2 = \frac{1}{2}(a + c) + \frac{1}{2}(a - c)\cos 2\Phi + \frac{1}{2}b\sin 2\Phi \quad (15)$$

onde $a = \mu_{20}$, $b = 2\mu_{11}$ e $c = \mu_{02}$.

Diferenciando-se a equação (15), igualando o seu resultado a zero e resolvendo para Φ , encontramos:

$$\tan 2\Phi = \frac{b}{a - c} \quad (16)$$

$$\sin 2\Phi = \pm \frac{b}{\sqrt{b^2 + (a - c)^2}} \quad (17)$$

$$\cos 2\Phi = \pm \frac{a - c}{\sqrt{b^2 + (a - c)^2}} \quad (18)$$

O eixo de orientação é obtido para o valor mínimo de X^2 . Se $b=0$ e $a=c$ o objeto não possui um único eixo de orientação. A alongação E do objeto é a razão entre o valor mais alto de X e o seu valor mais baixo, conforme mostra a equação (19). O valor mínimo e máximo de X^2 é determinado pelos sinais de $\sin 2\Phi$ e $\cos 2\Phi$ quando substituídos na equação (15). A alongação é mínima e vale 1 para os círculos.

$$E = \frac{X_{\max}}{X_{\min}} \quad (19)$$

Após determinar o valor de Φ que representa a orientação que o eixo principal faz com o eixo x , o próximo passo é rotacionar a figura com um ângulo Φ no sentido horário a fim de levar seu eixo principal para a posição horizontal. O efeito dessa rotação nos momentos é expresso pela equação (20)

$$M_{\Phi}(p, q) = \sum_{r=0}^p \sum_{s=0}^q (-1)^{q-s} \binom{p}{r} \binom{q}{s} (\cos \Phi)^{p-r+s} (\sin \Phi)^{q-s+r} M(p+q-r-s, r+s) \quad (20)$$

A última propriedade a ser verificada é se a figura possui uma inclinação positiva ao longo do eixo x . Para isso, precisa-se checar se o sinal de μ_{30} é positivo, caso contrário uma rotação adicional de 180° é necessária.

Existem muitas outras normalizações que são menos comuns mas que podem ser extremamente úteis, por exemplo pode-se citar a normalização em relação à reflexão e à relação de aspecto, geralmente aplicada após as normalizações de translação, escala e rotação.

Um conjunto de *momentos padrões* (*standard moments*), normalizado em relação a escala, translação, rotação e opcionalmente por reflexão, possui os seguintes valores:

- $M_{00}=1$ (área normalizada como 1)
- $M_{10}=0$ e $M_{01}=0$ (momentos centrais tomados sobre o centro de gravidade)
- $M_{11}=0$ (normalização da rotação $\theta=0$; momentos são tomados sobre o eixo principal alinhado em x).

Também pela normalização da rotação tem-se:

- $M_{20} \geq M_{02}$ e $M_{30} \geq 0$. Essa restrição serve para forçar ϕ a ser único, se as restrições não são alcançadas adiciona-se $\pi/2$ até que sejam satisfeitas. Se for selecionada normalização de reflexão:

- $M_{03} \geq 0$

Os *momentos padrões* representam características muito convenientes para análise de forma e pode ser comparado diretamente com outro conjunto de momentos padrões como uma medida de similaridade. Os *momentos padrões* podem ser considerados como um conjunto de momentos invariantes. Tal conjunto deve conter toda a informação dos momentos originais exceto para o ângulo de orientação do segmento de

imagem. Perceba que as informações dos momentos m_{00} , m_{10} , m_{01} e m_{11} são sacrificadas a fim de determinar as normalizações de massa, posição e rotação.

No método descrito acima, devido a propriedade de independência rotacional, os numerais “6” e “9” não podem ser distinguidos. O uso dos momentos de terceira ordem nesse caso é essencial.

Reeves et al., (1988) apresentaram um procedimento para análise tridimensional da forma a partir de imagens bidimensionais vista em um ângulo arbitrário. Duas técnicas globais foram testadas: Momentos e Descritores de Fourier, mas o trabalho focaliza especialmente nos *momentos padrões*. Os Descritores de Fourier (DF) aplicados sobre o contorno e Momentos aplicados sobre as silhuetas e contornos são comparados. Os momentos padrões obtiveram um resultado um pouco melhor que os descritores de Fourier e ambas as técnicas apresentaram resultados melhores que os invariantes de momento. Através da normalização da razão de aspecto da imagem, Reeves e seus colaboradores apresentaram um esquema eficiente para a identificação de objetos tridimensionais (seis aviões) vistos de um ângulo arbitrário. Aplicações em Visão por robô também são possíveis, especialmente no caso onde os pontos de vistas podem ser arbitrários e reduzidos.

Dos $\frac{1}{2}(n^2 + 3n + 2)$ momentos possíveis, considerando os quatro valores dos momentos padrões como constantes, existem $\frac{1}{2}(n^2 + 3n - 6)$ invariantes de momento. Hu definiu sete momentos invariantes de ordem 3, seis deles são invariantes absolutos e o sétimo momento é pseudo-invariante. Os invariantes absolutos são independentes entre si e o pseudo-invariante muda seu sinal sob reflexão, mas sua magnitude não é independente das outras seis. Quando a reflexão não é importante os pseudo-invariantes podem ser omitidos e alternativamente quando a reflexão é importante, um dos seis invariantes absolutos pode ser omitido do conjunto. Todos os valores devem ser usados apenas quando a importância da reflexão da imagem não é conhecida e um classificador simples é utilizado como por exemplo um classificador discriminante linear.

Baslev (1998) estudou a tolerância à ruídos dos momentos invariantes à rotação e à escala. O ruído considerado é aquele causado pela distorção de discretização através do uso de uma resolução espacial finita. As contribuições mais importantes de ruído são os provenientes do equipamento de digitalização (*scanner*, câmera de vídeo, *frame grabber*) e o ruído causado pelo arredondamento do pixel na discretização espacial. O ruído do

detector pode estar presente em qualquer localização da imagem enquanto que o ruído de discretização distorce principalmente as bordas e os componentes da imagem. Nesse trabalho, preocupou-se apenas com os ruídos causados pela discretização espacial por serem mais críticos e dominantes, exemplos desse ruído estão ilustrados na figura A2.4. O método utilizado para investigação por Baslev foi o dos Eixos Principais, onde os momentos invariantes à rotação são definidos através da transformação dos momentos de em torno do centro de gravidade.

Figura A2.4 – Caracter “a” da fonte “New Courier” amostrado em alta resolução (canto superior à esquerda) e amostras em resoluções mais baixas e rotacionadas em quatro diferentes ângulos (Baslev, 1998).

O trabalho de Baslev procurou responder as seguintes questões:

- Quão sensível é a invariância à rotação em relação à discretização espacial.
- Quanto essa sensibilidade depende da razão área/perímetro dos componentes da imagem
- Quanto essa sensibilidade depende da ordem dos momentos
- Se diferentes mapeamentos dos momentos dentro de espaços de características no reconhecimento de padrões clássico fornecem mesmos resultados
- Como a combinação de transformações de rotação e escala é comparada com a transformação de rotação apenas.

7 - INVARIÂNCIA À REFLEXÃO

Os momentos pares são naturalmente invariantes à reflexão. Para se conseguir a invariância à reflexão basta verificar se μ_{03} é não negativo. Se for, inverte-se a figura em

relação ao eixo X resultando que todos os momentos que dependem das potências ímpares de Y mudem de sinal. Isso pode ser expresso pela equação 21:

$$\dot{\mu}_{pq} = (-1)^q \mu_{pq} \quad (21)$$

Em algumas aplicações, a normalização da reflexão pode ser útil, nesse caso uma imagem e sua imagem espelho devem ser representadas pelo mesmo conjunto de momentos normalizados.

8 – INVARIÂNCIA À EXPANSÃO, COMPRESSÃO E INCLINAÇÃO

Alt, (1962) utilizou as variâncias (σ_x, σ_y) definidas pelas equação (23) para normalizar os momentos antes de normalizá-los em relação à área e com isso conseguiu invariância à expansão (*stretching*) and compressão (*squeezing*) nas direções x e y. Matematicamente, essas transformações, conhecidas como *affine transformations*, preservam as direções x e y e são da forma mostrada pela equação (22). O caso especial quando $a=c=1$ corresponde à translação. A equação (23) mostra os respectivos momentos, normalizados primeiro em relação à translação, a variância (desvio padrão) é utilizada para normalizar as coordenadas e em seguida os momentos são normalizados em relação à área.

$$\begin{aligned} X^* &= aX + b \\ Y^* &= cY + d \end{aligned} \quad (22)$$

$$\begin{aligned} \mu_{pq}^* &= \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} \frac{(x^*)^p (y^*)^q f(x, y)}{\mu_{00}} \\ x^* &= \frac{(x - \bar{x})}{\sigma_x} \\ y^* &= \frac{(y - \bar{y})}{\sigma_y} \\ \sigma_x &= \sqrt{\mu_{20} / \mu_{00}} \\ \sigma_y &= \sqrt{\mu_{02} / \mu_{00}} \end{aligned} \quad (23)$$

Um outro tipo de *affine transformations* derivadas de momentos, utilizada por Alt, (1962) foi feita fixando-se os pontos do eixo x. Isso permite que ocorra uma pequena variação (*slant*) dos padrões preservando a direção x, isto é, transformações do tipo mostrado pela equação (24). Um exemplo de aplicação dessa transformação seria permitir a pequena inclinação dos caracteres itálicos enquanto que omite a rotação de ângulos muito grandes que impossibilitaria por exemplo a distinção entre os caracteres “6” e “9”.

$$\begin{aligned} X^* &= aX + eY + b \\ Y^* &= cY + d \end{aligned} \quad (24)$$

Para tornar os momentos invariantes à esse tipo de transformação, Alt fez uma nova normalização, conforme ilustra a equação 25

$$\begin{aligned} \rho &= \frac{\sum_{x=0}^{M-1} \sum_{y=0}^{N-1} (x^*) (y^*) f(x, y)}{\sum_{x=0}^{M-1} \sum_{y=0}^{N-1} (y^*)^2 f(x, y)} \\ x &= \frac{x^* - \rho \cdot y^*}{\sqrt{1 - \rho^2}} \\ y &= y^* \end{aligned} \quad (25)$$

$$\mu_{pq} = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} x^p y^q f(x, y)$$

Os momentos definidos pela equação (25), após normalização apresentam em particular $\ell_{00}=\ell_{20}=\ell_{02}=1$ e $\ell_{10}=\ell_{01}=\ell_{11}=0$. A quantidade ρ é chamada em Estatística de coeficiente de regressão de x^* em y^* .

Raveendran et al., (1993) apresentaram uma técnica para classificar imagens que tenham sido alongadas ou contraídas. Segundo mostram os autores, a imagem não permanece invariante quando é escalonada apenas em uma direção (x ou y). Os invariantes propostos são invariantes à reflexão, translação e escala mas não à rotação. Os autores demonstraram que os momentos convencionais só são invariantes quando a

$$\text{aspecto} = 2 \log \left(\frac{\frac{v_{\max}}{v_{\min}} - 0.99999}{M - 1} \right)$$

Nas equações acima V é a matriz de covariância e $v_{\max}$ e $v_{\min}$ são os autovalores da matriz. A constante M é aproximadamente igual a média da razão $\frac{v_{\max}}{v_{\min}}$. O valor 2

imagem é escalonada igualmente duas direções, conforme mostra a equação (26), onde a representa o fator de escala na direção x e b representa o fator de escala na direção y . Os momentos só serão invariantes apenas $a=b$.

$$\gamma_{pq} = \left(\frac{b}{a}\right)^{\frac{p-q}{2}} \mu_{pq} \quad (26)$$

Quando $a \cong b$, Raveendran et al. (1993) utilizaram os seguintes invariantes:

$$R_{pq} = \gamma_{pq} \gamma_{qp} = \left(\frac{b}{a}\right)^{\frac{p-q}{2}} \eta_{pq} \left(\frac{b}{a}\right)^{\frac{q-p}{2}} \eta_{qp} = \eta_{pq} \eta_{qp} \quad (27)$$

9 – INVARIÂNCIA À RAZÃO DE ASPECTO

Reeves et al., (1988) desenvolveu a normalização para razão de aspecto, a fim de evitar que em objetos muito alongados os momentos na dimensão principal assumam valores muito grandes que possam desbalancear o peso de outras características importantes da imagem. Nesses casos o ideal seria que a razão entre os eixos principais da figura fosse igual a 1. Isso pode ser feito comprimindo o objeto ao longo de sua dimensão principal, que é rotacionada para ser a direção x .

O processo de compressão horizontal do objeto consiste em aumentar a sua altura, enquanto mantém sua área constante, para isso basta multiplicar a coordenada x por uma constante, enquanto divide-se a coordenada y pela mesma constante. Quando uma figura é expandida em uma direção particular, multiplicando-se a mesma por uma constante, os momentos daquela direção são multiplicados pela constante elevada pela potência de 1 mais a ordem do momento naquela direção. Momentos na direção perpendicular à direção expandida não são afetados. A Equação 28 expressa o efeito nos momentos causado por essas duas mudanças recíprocas.

$$\mu_{pq}^a = a^{p+1} \left(\frac{1}{a}\right)^{q+1} \mu_{pq} \quad (28)$$

A condição para que o espalhamento em ambas as direções seja igual é mostrado na equação 29 abaixo:

$$\mu_{20}^a = \mu_{02}^o \quad (29)$$

Substituindo a equação 28 em 29 e resolvendo obtém-se:

$$a = \left(\frac{\mu_{02}}{\mu_{20}} \right)^{\frac{1}{4}} \quad (30)$$

A equação (30) pode ser usada com a equação (28) para computar todos os momentos após a normalização para razão de aspecto.

Em muitas aplicações a razão de aspecto da figura original é uma característica importante da forma e a normalização de aspecto é descartada. Seu valor pode assumir valores entre 0 e 1.

Master, (1994) sugere uma outra transformação para calcular a razão de aspecto de uma figura, conforme mostra as equações (31) a seguir:

$$\begin{aligned} V &= \begin{bmatrix} \mu_{20} & \mu_{11} \\ \mu_{11} & \mu_{02} \end{bmatrix} \\ a &= \mu_{20} \\ b &= \mu_{11} \\ c &= \mu_{02} \\ r &= \sqrt{(a - c)^2 + 4b^2} \\ v_{\max} &= \frac{1}{2}(a + c + r) \\ v_{\min} &= \frac{1}{2}(a + c - r) \\ \text{aspecto} &= 2 \log \left(\frac{\frac{v_{\max}}{v_{\min}} - 0.99999}{M - 1} \right) \end{aligned} \quad (31)$$

Nas equações acima V é a matriz de covariância e $v_{\max}$ e $v_{\min}$ são os autovalores da matriz. A constante M é aproximadamente igual a média da razão $\frac{v_{\max}}{v_{\min}}$. O valor 2

que multiplica a expressão pode ser modificado de maneira a produzir uma faixa adequada para os valores transformados. O autor sugere que $M=2$ pode acomodar um

grande número de formas.

Uma outra característica interessante que pode ser utilizada em aplicações onde não ocorra orientações desconhecidas e aleatórias, é a razão de aspecto fixa mostrada na equação 32. Ela é particularmente sensível à altura de uma figura com respeito à sua largura e é útil em distinguir por exemplo árvores de trens.

$$af = k_1 \log\left(\frac{\mu_{02}}{\mu_{20}}\right) - k_2 \quad (32)$$

Master, (1994) sugere $k_1=2$ e $k_2=0$ onde uma grande faixa de figuras podem ser acomodadas.

Em relação às medidas relativas dos comprimentos dos eixos principais, a medida chamada de espalhamento $\mu_{02}+\mu_{20}$ vale $1/2\pi$ para um círculo perfeito, considerando a normalização em relação à área, onde $\mu_{00}=1$. Todas as outras formas apresentarão valores maiores para essa soma. Objetos muito alongados produzem grandes somas. Alguns aspectos dessa soma são interessantes, como por exemplo sua invariância à rotação. Ela é também igual a soma dos autovalores mostrada na equação (31).

Uma transformação geralmente eficiente para essa variável é mostrada na equação (33):

$$spread' = K_1 \log\left(\frac{v_{max} + v_{min} - 0.5/\pi + \varepsilon}{K_2 - 0.5/\pi + \varepsilon}\right) \quad (33)$$

onde K_2 é determinado através da soma média de um conjunto representativo de figuras. Isso força a transformação a mapear o valor média para zero. Ignorando ε calcula-se K_1 fazendo um mapeamento do valor máximo da soma para o valor 10. Depois calcula-se e considerando um mapeamento do valor mínimo $1/2\pi$ para um valor negativo grande. Master, (1994) sugeriu os seguintes valores $K_1=10$, $K_2=0.25$ e $\varepsilon=0.001$.

10 – INVARIÂNCIA À CONTRASTE

Maitra (1979) apresentou uma variação dos invariantes de Hu com um invariante adicional à mudança de contraste e demonstrou a invariância de duas imagens digitalizadas de uma mesma cena que foram tomadas com uma câmera em posições diferentes a fim de fornecer uma diferença de escala, iluminação, posição e rotação. Em relação aos momentos de Hu, os momentos de Maitra são dados pela equação (34) a seguir:

$$\begin{aligned}
 \beta_1 &= \frac{\sqrt{\phi_2}}{\phi_1} \\
 \beta_2 &= \frac{\phi_3 \mu_{00}}{\phi_1 \phi_2} \\
 \beta_3 &= \frac{\phi_4}{\phi_3} \\
 \beta_4 &= \frac{\sqrt{\phi_5}}{\phi_4} \\
 \beta_5 &= \frac{\phi_6}{\phi_1 \phi_4} \\
 \beta_6 &= \frac{\phi_4}{\phi_3}
 \end{aligned} \quad (34)$$

Abo-Zaid et al. (1988) também propuseram uma variação dos momentos de Hu para correção das mudanças de escala e contraste antes do cálculos dos invariantes. Em termos dos momentos centrais o novo fator de normalização é definido por

$$\mu'_{pq} = \mu_{pq} \frac{1}{\mu_{00}} \left(\frac{\mu_{00}}{\mu_{20} + \mu_{02}} \right)^{(p+q)/2} \quad (35)$$

11 – INVARIÂNCIA À FALTA DE NITIDEZ NA IMAGEM (*BLURING*)

Flusser et al., (1996) apresentaram um conjunto de invariantes simétricos à perda de nitidez (*blurring*) em imagens, que ocorre devido às condições imperfeitas no processo de formação da imagem tais como difração, aberração das lentes, foco errado e turbulência atmosférica. Na derivação dos invariantes, os autores consideraram as condições de simetria dada por $\ell_{pq} = \ell_{qp}$ e se p e/ou q são ímpares, então $\ell_{pq} = 0$. Os

invariantes à falta de nitidez são listados de quarta e quinta ordem são mostrados na equação (36) a seguir.

$$\begin{aligned}
 B(1,3) &= \mu_{13} - \frac{3\mu_{02}\mu_{11}}{\mu_{00}} \\
 B(3,1) &= \mu_{31} - \frac{3\mu_{20}\mu_{11}}{\mu_{00}} \\
 B(4,0) &= \mu_{40} - \mu_{04} - \frac{6\mu_{02}(\mu_{20} - \mu_{02})}{\mu_{00}} \\
 B(3,2) &= \mu_{32} - \frac{3\mu_{12}\mu_{20} + \mu_{30}\mu_{02}}{\mu_{00}} \\
 B(2,3) &= \mu_{23} - \frac{3\mu_{21}\mu_{02} + \mu_{03}\mu_{20}}{\mu_{00}} \\
 B(4,1) &= \mu_{41} - \frac{6\mu_{21}\mu_{20}}{\mu_{00}} \\
 B(0,5) &= \mu_{05} - \frac{10\mu_{03}\mu_{02}}{\mu_{00}} \\
 B(5,0) &= \mu_{50} - \frac{10\mu_{30}\mu_{20}}{\mu_{00}}
 \end{aligned} \tag{36}$$

Os invariantes devem ser ainda normalizados de maneira a ser independente do contraste da imagem e de maneira a ter o mesmo “peso” no espaço euclidiano, conforme a equação (37), onde N é o tamanho da imagem.

$$B'(p,q) = \frac{B(p,q)}{\mu_{00} \cdot (N/2)^{p+q}} \tag{37}$$

Também é possível utilizar a normalização abaixo, onde as características são invariantes à escala mas sensíveis à contraste.

$$B''(p,q) = \frac{B(p,q)}{\mu_{00}^{(p+q+2)/2}} \tag{38}$$

12 - INTERPRETAÇÃO GEOMÉTRICA DOS MOMENTOS CARTESIANOS

O uso de Momentos como descritor global de características é conveniente do ponto de vista de processamento de imagem uma vez que só requer poder computacional, o que é abundante atualmente, entretanto o grande problema refere-se à seleção do conjunto de características e também em como interpretar a informação fornecida pelos momentos.

Vários atributos de formato podem ser extraídos dos Momentos Cartesianos. Conforme visto na seção anterior, o momento de ordem zero m_{00} fornece uma medida de área do objeto. Com os momentos de primeira ordem m_{10} e m_{01} , denominados “momentos de alavanca”, calcula-se o centro de massa do objeto, que é utilizado como uma localização padrão do objeto dentro do campo visual. As coordenadas do centro de massa é a interseção das linhas $x = \bar{x}$ e $y = \bar{y}$. Essas linhas são paralelas aos eixos x e y e representam as linhas onde toda a massa pode estar concentrada sem alterar os primeiros momentos sobre os eixos x e y . O centro de massa define uma localização única do objeto e pode ser usado como ponto de referência para descrever a posição do objeto dentro do campo de visão. Se a posição de um objeto é coincidente com a origem do campo visual ou seja $\bar{x} = 0$ e $\bar{y} = 0$, então os momentos computados para aquele objeto são denominados *Momentos Centrais* e são designados por ℓ_{pq} (note que $\ell_{10} = \ell_{01} = 0$). Em algumas aplicações industriais a posição é dada pelo retângulo que envolve o objeto, entretanto o centro de massa representa um ponto relativamente insensível à ruído na imagem.

Os momentos m_{10} e m_{01} podem ser usados também para calcular o ângulo T do eixo de mínima inércia ou eixo principal da imagem (Strachan et al. 1990):

$$T = \frac{1}{2} \arctan \frac{2(m_{00}m_{11} - m_{10}m_{01})}{(m_{00}m_{20} - m_{10}^2) - (m_{00}m_{02} - m_{01}^2)} \quad (40)$$

Os momentos de segunda ordem m_{02} , m_{11} e m_{20} são conhecidos como *Momentos de Inércia* e determinam várias características interessantes tais como os eixos principais sobre os quais os segundos momentos são mínimo e máximos (eixo maior e eixo menor respectivamente), a elongação e orientação do eixo principal do objeto. Os momentos m_{20} e m_{02} são denominados “momentos de inércia” e o momento m_{11} é denominado

produto de inércia. Os momentos de segunda e terceira ordem conduzem a uma série de interpretações intuitivas que carregam uma natureza Gestalt (Giuliano et al., 1961).

O ângulo Φ que o eixo principal maior ou menor (o mais próximo) faz com o eixo é mostrado na equação 41 a seguir. A orientação do *eixo principal maior* θ pode ser especificamente determinada a partir dos valores específicos de μ_{11} , $(\mu_{20}-\mu_{02})$ e Φ , como ilustra a tabela A2.2 extraída de Prokop & Reeves, (1992). O ângulo θ sozinho não garante uma única orientação uma vez que a ambigüidade de 180° ainda existe. Os momentos de terceira ordem podem ser usados para resolver essa ambigüidade.

$$\Phi = \frac{1}{2} \tan^{-1} \left(\frac{2\mu_{11}}{\mu_{20} - \mu_{02}} \right) \quad (41)$$

Tabela A2.2 – Orientação do eixo principal

μ_{11}	$\mu_{20}-\mu_{02}$	Φ	θ
0	-	0	$\pi/2$
+	-	$0 > \phi > -\pi/4$	$\pi/2 < \phi < \pi/4$
+	0	0	$\pi/4$
+	+	$\pi/4 > \phi > 0$	$\pi/4 > \phi > 0$
0	0	0	0
-	+	$0 > \phi > -\pi/4$	$0 > \phi > -\pi/4$
-	0	0	$-\pi/4$
-	-	$\pi/4 > \phi > 0$	$-\pi/2 > \phi > -\pi/4$

O ângulo θ do eixo principal de menor inércia pode ser usado como ponto de referência único para descrever a orientação do objeto dentro do campo (plano de rotação).

As interpretações dos vários momentos de baixa ordem mostradas na tabela A2.3 possui uma analogia com a mecânica dos corpos e foram descritas em Giuliano et al, (1961).

Tabela A2.3 – Interpretação geométrica dos Momentos Centrais de segunda e terceira Ordem

Momentos Centrais	Interpretação Geométrica
μ_{20}	Centralidade horizontal
μ_{02}	Centralidade vertical
μ_{11}	Diagonalidade
μ_{12}	Divergência horizontal
μ_{21}	Divergência vertical
μ_{30}	Desequilíbrio horizontal
μ_{03}	Desequilíbrio vertical

A interpretação é feita com o ponto de vista centralizado sobre o centro de gravidade de cada objeto, com os eixos das coordenadas colocados sobre ele. A figura A2.5 extraída de Giuliano et al. (1961) mostra a relação entre o centro de gravidade e o centro geométrico de várias letras, conforme exemplifica as definições dos momentos a seguir.

- A *centralidade horizontal* μ_{20} é determinada pela variância da massa na direção x. Em uma imagem do objeto ela mede a variância da forma do objeto na direção x.

- A *centralidade vertical* μ_{02} é determinada pela variância da massa na direção y. Em uma imagem do objeto ela mede a variância da forma do objeto na direção y.

- A *diagonalidade* μ_{11} é a distribuição da massa sobre os quatro quadrantes. Ela mede a proporcionalidade da distribuição da forma do objeto entre os 1º e 3º e os 2º e 4º quadrantes, vista do centro de massa.

- Se $\mu_{11} < 0$ há mais massa nos 2º e 4º quadrantes (letra “L”).
- Se $\mu_{11} > 0$ há mais massa nos 1º e 3º quadrantes (letras “Z”, “P”, “J”)
- Se $\mu_{11} = 0$ há uma distribuição de massa igual nos 1º e 3º em relação ao 2º e 4º quadrantes (letras “A”, “E”, “O”)

- A *divergência horizontal* μ_{12} mede a dimensão da figura à esquerda comparada com a sua direita, ou seja, a simetria do objeto na direção x.

- Se $\mu_{12} < 0$ há mais massa nos 2º e 3º quadrantes (letras “P”, “L”, “E”).
- Se $\mu_{12} > 0$ há mais massa nos 1º e 4º quadrantes (letras “J”)

- Se $\mu_{12} = 0$ há uma distribuição de massa igual nos 2º e 3º em relação ao 1º e 4º quadrantes (letras “A”, “Z”, “O”)

• A *divergência vertical* μ_{21} mede a dimensão da parte inferior da figura (*bottom*) comparada com a sua parte superior (*top*), ou seja, a simetria do objeto na direção y.

- Se $\mu_{21} < 0$ há mais massa nos 3º e 4º quadrantes (letras “A”, “L”, “J”).
- Se $\mu_{21} > 0$ há mais massa nos 1º e 2º quadrantes (letra “P”)
- Se $\mu_{21} = 0$ há uma distribuição de massa igual nos 3º e 4º em relação ao 1º e 2º quadrantes (letras “Z”, “E”, “O”)

• O *desequilíbrio horizontal* μ_{30} determina a localização do centro de gravidade CG com relação à metade da largura ($X_{l/2}$) do objeto.

- Se $\mu_{30} < 0$ CG à direita de $X_{l/2}$ (letra “J”).
- Se $\mu_{30} > 0$ CG à esquerda de $X_{l/2}$ (letras “P”, “E” e “L”).
- Se $\mu_{30} = 0$ $\bar{x} = X_{l/2}$ (letras “A”, “Z”, “O”)

• O *desequilíbrio vertical* μ_{03} determina a localização do centro de gravidade com relação a metade da altura ($Y_{h/2}$) do objeto.

- Se $\mu_{03} < 0$ CG acima de $Y_{h/2}$ (letras “P”, “E”).
- Se $\mu_{03} > 0$ CG abaixo de $Y_{h/2}$ (letras “A”, “L” e “J”).
- Se $\mu_{03} = 0$ $\bar{y} = Y_{h/2}$ (letras “Z”, “O”)

Figura A2.5 – Relação entre centro de gravidade e centro geométrico de algumas letras (Giuliano et al., 1961)

A tabela A2.4 extraída de Giuliano et al. (1961) sumariza os sinais obtidos para as letras mostradas na figura A2.5.

Tabela A2.4 – Sinais obtidos para os Momentos Centrais por Giuliano et al. (1961)

Letras	μ_{03}	μ_{30}	μ_{21}	μ_{12}	μ_{11}
A	+	0	–	0	0
P	–	+	+	–	+
L	+	+	–	–	–
E	–	+	0	–	0
J	+	–	–	+	+
O	0	0	0	0	0
Z	0	0	0	0	+

As variâncias μ_{20} e μ_{02} são medidas de dispersão estatística que caracteriza a variabilidade de uma função densidade e mede as variações de amplitude da forma nas direções x e y.

Os momentos de segunda ordem também definem uma aproximação inercialmente equivalente à imagem original referida como *ellipse da imagem* (Teague, 1980). A *ellipse da imagem* é um disco elíptico de intensidade constante, que possui a mesma massa e os mesmos momentos de segunda ordem da imagem original. A imagem da *ellipse* pode ser definida pelo semi-eixo maior α ao longo do eixo x e semi-eixo menor β ao longo do eixo y, conforme equações (42) e (43) abaixo:

$$\alpha = \left(\frac{2 \left[\mu_{20} + \mu_{02} + \sqrt{(\mu_{20} - \mu_{02})^2 + 4\mu_{11}^2} \right]}{\mu_{00}} \right)^{\frac{1}{2}} \tag{42}$$

$$\beta = \left(\frac{2 \left[\mu_{20} + \mu_{02} - \sqrt{(\mu_{20} - \mu_{02})^2 + 4\mu_{11}^2} \right]}{\mu_{00}} \right)^{\frac{1}{2}} \tag{43}$$

A figura A2.6 extraída de Prokop & Reeves, (1992) ilustra a elipse da imagem para a silhueta de um objeto. A fim de que a elipse tenha os mesmos momentos de segunda ordem da imagem original é necessário que a elipse esteja centrada no centro de massa do objeto e rotacioná-la com um ângulo Φ , a fim de alinhar seu eixo maior com o eixo principal da figura.

Figura A2.6 – Elipse da Imagem

A intensidade da imagem da elipse é dada pela equação (44)

$$I = \frac{\mu_{00}}{\pi \alpha \beta} \quad (44)$$

Outra propriedade que pode ser determinada pelos momentos de segunda ordem é o Raio de Giração (RG) da imagem. O Raio de Giração⁷ sobre um eixo é a distância do eixo à linha onde toda a massa está concentrada, sem alterar o segundo momento sobre o eixo. Os Raios de Giração sobre os eixos x e y são respectivamente dados pelas equações (45) e (46) abaixo:

$$RG_x = \sigma_x = \sqrt{\frac{\mu_{20}}{m_{00}}} \quad (45)$$

⁷ O Raio de Giração pode ser interpretado também como o raio mínimo que inclui a maioria senão todos os pontos em seu interior.

$$RG_y = \sigma_y = \sqrt{\frac{\mu_{02}}{m_{00}}} \quad (46)$$

Os Raios de Giração RG_x e RG_y medem respectivamente o desvio padrão σ_x e σ_y do objeto nas direções x e y .

O Raio de Giração sobre a origem é o raio de um círculo centrado na origem onde toda a massa está centrada sem alterar o segundo momento sobre a origem e é dado pela equação 47 a seguir:

$$RG_{xy} = \sqrt{\frac{\mu_{20} + \mu_{02}}{\mu_{00}}} \quad (47)$$

O Raio de Giração em relação à origem é inerentemente invariante à orientação da imagem e consequentemente tem sido utilizado como característica invariante à rotação para a representação de objetos.

Belkasim et al., (1991) determinou a invariância à escala dividindo os momentos pelo fator $(M_{20} + M_{02})^z$ onde $z = (p+q+2)/4$.

Considerando um sistema de coordenadas fixas XYZ adaptadas a um modelo de aviões em particular, Dudani et al., (1977) normalizou os momentos em relação a variação de tamanho, através do Raio de Giração, tornou-os invariantes em relação ao ângulo de elevação θ (em relação ao eixo Y) e à distância B ao longo do eixo óptico de uma câmera⁸, conforme mostra a equação 48

⁸ O Raio de Giração é diretamente proporcional ao tamanho da imagem ou inversamente proporcional à distância do objeto ao longo do eixo óptico (B), assim o produto do Raio de Giração da imagem e a distância ao longo do eixo óptico do objeto é uma constante e pode ser usado para determinar invariância ao tamanho, ao ângulo de elevação e à distância ao longo do eixo óptico.

$$\begin{aligned}
\phi'_1 &= (\mu_{20} + \mu_{02})^{1/2} \cdot B = r \cdot B \\
\phi'_2 &= \phi_2 / r^4 \\
\phi'_3 &= \phi_3 / r^6 \\
\phi'_4 &= \phi_4 / r^6 \\
\phi'_5 &= \phi_5 / r^{12} \\
\phi'_6 &= \phi_6 / r^8 \\
\phi'_7 &= \phi_7 / r^{12}
\end{aligned} \tag{48}$$

Essas características de momento podem ser usadas para representar ou caracterizar a imagem para uma dada visão de aspecto do objeto, que é definido unicamente pelo ângulo de azimute (em relação ao eixo Z) e o ângulo de rotação (*roll*) em relação ao eixo X. Nesse trabalho Dudani e seus colegas extraíram dois conjuntos de momentos, um para as bordas e um para a silhueta dos aviões. Segundo eles, os momentos derivados dos contornos contêm mais informações sobre partes de frequência alta da imagem do que os momentos derivados das silhuetas, assim os detalhes tais como formato do nariz do avião ou sua cauda, são melhores caracterizados pelos momentos computados a partir dos contornos. Por outro lado, as características estruturais mais grosseiras do avião tais como a fuselagem e as asas são melhores caracterizadas pelas silhuetas e menos suscetíveis á ruídos, tem-se assim um vetor de características de dimensão 14 para a imagem de um objeto para qualquer valor de azimute e de rotação. Um método é aplicado para se determinar as características mais significativas e apenas um subconjunto das cinco características mais informativas é selecionado. Na classificação das imagens (seis tipos) utilizou-se a regra de decisão de Bayes e a regra do vizinho mais próximo, que apresentou um resultado mais satisfatório.

Outras características comumente utilizadas no Reconhecimento de Padrões são o espalhamento (*spread*) dado pelo invariante de momento $\phi_1 = \mu_{20} + \mu_{02}$ e que de maneira análoga à μ_{00} fornece uma informação de tamanho do objeto e a delgadeza (*slenderness*) do objeto dada pela equação (49).

$$\sqrt{\phi_2} = \sqrt{(\mu_{20} - \mu_{02})^2 + 4\mu_{11}^2} \tag{49}$$

Ballard & Brown (1982) sugerem como medida de excentricidade e a seguinte equação 50:

$$e = \frac{(\mu_{20} - \mu_{02})^{1/2} + 4\mu_{11}}{m_{00}} \quad (50)$$

Otsu (1986) definiu um descritor de forma “invariante à similaridade” definido pela equação (51) abaixo:

$$S = \frac{\sqrt{4\mu_{11}^2 + (\mu_{20} - \mu_{02})^2}}{\mu_{20} + \mu_{02}} \quad (51)$$

onde $0 \leq S \leq 1$. O valor máximo $S=1$ é obtido se o padrão for uma linha reta e é mínimo se não for correlacionado ($\ell_{11} = 0$) e isotrópico ($\ell_{20} = \ell_{02}$).

Hu (1962) simulou um simples modelo de reconhecimento de padrões usando como características os dois primeiros invariantes de momento. O ponto $X=\phi_1$ e $Y=\sqrt{\phi_2}$ é usado para representar um padrão em um espaço bidimensional. Considerando os padrões aprendidos anteriormente e representados por (X_i, Y_i) , $i=1,2,\dots,n$ dentro de suas respectivas classes. Para cada novo padrão (X,Y) apresentado ao modelo, calcula-se a distância entre (X,Y) e (X_i, Y_i) dada pela equação (52) abaixo:

$$d_i = \sqrt{(X - X_i)^2 + (Y - Y_i)^2} \quad (52)$$

A distância mínima (melhor distância) pode ser definida como $d_{\min} = \min (d_i)$. Essa distância é então comparada com um *threshold* T e se $d_{\min} > T$, o programa pede que se rotule o novo padrão e armazena (X,Y) como (X_{n+1}, Y_{n+1}) , assim um novo padrão pode ser aprendido.

Os momentos do tipo m_{p0} e m_{0q} são equivalentes aos *momentos da projeção* da imagem nos eixos x e y respectivamente, conforme ilustra a figura A2.7 extraída de Prokop & Reeves, (1992). A projeção vertical $v(x)$ de uma imagem segmentada $f(x,y)$ dentro de um eixo x é dada pela equação (53) a seguir.

$$v(x) = \int_{-\infty}^{\infty} f(x,y) dy \quad (53)$$

Os momentos unidimensionais m_p de $v(x)$ são dados por:

$$m_p = \int_{-\infty}^{\infty} x^p v(x) dx \quad (54)$$

Substituindo (53) em (54), obtém-se:

$$m_p = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^p f(x,y) dx dy = m_{p0} \quad (55)$$

Se considerarmos a projeção de uma imagem em um eixo como uma distribuição de probabilidade, as propriedades dos momentos centrais descrevem características estatísticas clássicas, conforme exemplifica o segundo momento de uma projeção da imagem no eixo x que é proporcional à variância da distribuição, e é dada por :

$$\mu_{20} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^2 f(x,y) dx \quad (56)$$

Figura A2.7 – Momentos de projeção de um objeto (Prokop & Reeves, 1992)

Os momentos de primeira ordem de uma imagem também são iguais aos primeiros momentos de suas projeções, assim como a área, dessa maneira a posição do objeto pode ser calculada a partir das projeções dos primeiros momentos. Os momentos

de Segunda ordem podem ser calculados a partir da projeção diagonal de uma imagem, assim a orientação pode ser computada utilizando as projeções horizontais, verticais e diagonais. As projeções oferecem uma representação compacta e permite a aplicação de algoritmos rápidos, por exemplo podem ser geradas a partir do *run-length code* de uma imagem (Jain et al, 1995).

Os dois momentos de terceira ordem μ_{30} e μ_{03} descrevem a *skewness* (inclinação, deformação) das projeções da imagem. *Skewness* é uma medida estatística clássica que mede o grau de desvio de simetria de uma distribuição sobre a média. Os coeficientes de Inclinação para as projeções da imagem nos eixos x e y são dados pela equação (57) abaixo

$$\begin{aligned} Sk_x &= \frac{\mu_{30}}{\mu_{20}^{3/2}} \\ Sk_y &= \frac{\mu_{03}}{\mu_{02}^{3/2}} \end{aligned} \tag{57}$$

Os sinais dos coeficientes SK_x e SK_y indica qual lado do eixo a projeção está inclinada conforme mostra a Tabela A2.5. Note que se $SK_x=0$ ou $SK_y=0$ não garante que o objeto é simétrico (Prokop & Reeves, 1992).

Tabela A2.5 - Inclinação das Projeções em relação a SK_x e SK_y

SK_x	Projeção x inclinada	SK_y	Projeção y inclinada
+	à esquerda do eixo y	+	abaixo do eixo x
0	simétrica em relação ao eixo y	0	simétrica em relação ao eixo x
-	à direita do eixo y	-	acima do eixo x

A inclinação é positiva quando a maioria dos pontos (metade) da forma amostrada possui um valor menor que o valor médio (centróide).

Os momentos de terceira ordem também podem ser usados para calcular a orientação da imagem de um objeto (Yakali et al., 1998), conforme a equação (58) abaixo:

$$\Phi = \arctan2(\mu_{21} + \mu_{30}, \mu_{30} + \mu_{12}) \quad (58)$$

Uma outra propriedade dos momentos de terceira ordem é que eles formam um conjunto de momentos ortogonais absolutos.

Uma técnica simples mas poderosa para localizar objetos consiste na assinatura direcional de níveis de cinza. A assinatura direcional consiste em uma função que soma os valores de níveis de cinza dos pixels ao longo da linha normal à linha de referência. Hall et al., (1971) citaram o uso de assinaturas para segmentação de objetos tais como coração, pulmão, diafragma e clavícula. A informação de assinatura permite realizar um “zoom” em um objeto em particular. Após um objeto em particular haver sido localizado, medidas como tamanho e textura pode ser realizadas.

Os dois momentos centrais de quarta ordem m_{04} e m_{40} descrevem a curtose da projeção da imagem. Curtose (coeficiente de excesso) é uma medida clássica estatística que mede o pico de uma distribuição. O coeficiente de Curtose para a projeção de uma imagem nos eixos x e y é dado pela equação 59 a seguir

$$\begin{aligned} K_x &= \frac{\mu_{40}}{\mu_{20}^2} - 3 \\ K_y &= \frac{\mu_{04}}{\mu_{02}^2} - 3 \end{aligned} \quad (59)$$

Uma Curtose igual a zero é o valor para a distribuição Gaussiana,. Valores menores que zero indica um pico de distribuição mais achatado enquanto valores mais positivos indicam uma distribuição mais estreita e pontiaguda. Uma curtose positiva significa que a maioria das amostras estão concentradas em torno da média.

Yakali & Gröse, (1998) definiram os termos $\eta_x = \ell_{30} + \ell_{12}$ e $\eta_y = \ell_{03} + \ell_{21}$ como desequilíbrio ao longo do eixo X e Y respectivamente.

Os momentos de terceira ordem podem ser usados para resolver a ambigüidade à rotação de 180° uma vez que o sinal do coeficiente de inclinação muda em ambos os eixos. Se a imagem for rotacionada por um ângulo - θ e o eixo principal coincidir com o eixo x o sinal de μ_{30} pode ser usado para distinguir entre as duas orientações possíveis.

Reeves & Rostampour (1981) utilizaram momentos padrões para análise global de formas genéricas, tais como retângulo, elipse, losango e objeto côncavo. A curtose K_x da projeção do eixo principal (eixo x) foi calculada para cada uma das formas computadas a partir dos momentos padrões. Adicionalmente, foi calculada o comprimento e largura normalizados, para as três primeiras formas, como função dos momentos de Segunda ordem. Esses valores podem ser vistos na tabela A2.6 (Prokop & Reeves, 1992).

Tabela A2.6 – Curtose e Dimensões normalizadas para formas genéricas

Forma	K_x	Comprimento	Largura
Retângulo	-1.2	$\sqrt{12M_{20}}$	$1/\sqrt{12M_{20}}$
Elipse	-1.0	$\sqrt{16M_{20}}$	$1/\pi\sqrt{M_{20}}$
Losango	- 0.6	$\sqrt{24M_{20}}$	$1/\sqrt{6M_{20}}$
Côncavo	- 0.6		

Uma outra propriedade qualitativa dos momentos invariantes é os componentes da imagem são ponderados progressivamente à medida que suas localizações se afastam do centro de massa da imagem, assim podemos pensar que o kernel $x^p y^q$ dos momentos geométricos decresce à medida que x e y se aproximam de zero. Isso indica que os momentos *salientam* a informação contida próximas às bordas dos objetos. Além do mais, devido a natureza de variação lenta das funções base dos momentos de baixa ordem, os detalhes macroscópicos são relacionados resulttando em um efeito de média (Freeman & Saleh, 1988).

Muitas aplicações utilizam momento até a terceira ordem, e algumas aplicações que apresentam baixo ruído e alta definição de características utilizam momento até quarta ordem. Acima da quarta ordem, os momentos tornam-se mais suscetíveis à pequenas variações na forma. Embora momentos sejam ótimos descritores de formato,

eles não representam uma boa solução quando detalhes finos são necessários (Master, 1994).

Os momentos invariantes discriminam bem classes nas quais a forma ou o contorno do padrão de entrada contém a maior parte da informação necessária para separar as classes. Às vezes é necessário uma característica diferente para lidar com classes que não são distinguidas apenas pela forma. Momentos de ordens superiores são uma opção mas já foi visto que quanto maior a ordem, maior a degradação devido ao ruído. Uma opção sugerida por Freeman & Saleh (1987) é utilizar momentos de baixa ordem da transformada de Fourier para aumentar o conjunto de características. Tais momentos de Fourier estão menos suscetíveis à ruídos do que são os momentos de ordens mais altas, dessa maneira, tem-se uma abordagem dual que resulta em um conjunto de características independentes da tarefa (*task-independent*) mais útil do que cada conjunto em seu domínio apenas.

Além de fornecerem uma representação concisa das propriedades geométricas fundamentais da imagem, transformações geométricas básica também podem ser realizadas sobre as representações dos momentos de uma imagem. Essas transformações são mais facilmente determinadas no domínio dos momentos do que no domínio dos pixels da imagem original. As transformações de escala, translação, rotação, reflexão, contraste e convolução resultam em uma nova imagem e os momentos transformados podem ser calculados em função dos momentos da imagem original. A convolução de uma imagem $f(x,y)$ com um *kernel* discreto $w(i,j)$ de tamanho $N \times M$ pode ser considerado a soma de uma série de translações e escalas (Prokop & Reeves, 1992).

Em particular os invariantes de momentos geométricos são altamente dependentes da simetria das formas. Os momentos envolvendo as potências ímpares de p (q) são zero para formas simétricas sobre o eixo vertical (horizontal). Qualquer momento contendo uma potência ímpar de p e q é zero para formas que são simétricas em ambas as direções vertical e horizontal. Para tais formas os alguns momentos são iguais (zero) e não fornecem informações para separação das classes. Em relação ao método dos Eixos Principais a grande dificuldade é determinação do ângulo θ . Quando a figura é assimétrica é fácil determinar a direção do único eixo principal, entretanto as vezes os eixos principais não podem ser unicamente definidos, particularmente quando a imagem é

circular ou apresenta simetria rotacional (*n-fold*)⁹. Por outro lado, se a imagem é simétrica pode se reduzir a quantidade de espaço de armazenamento para representação da forma, comparando apenas a porção repetitiva do objeto ao invés de toda a forma. Se certas porções de simetria forem perdidas, a simetria pode fornecer pistas para recuperar a forma., além do mais a ordem de simetria n representa uma boa característica de forma uma vez que é invariante a escala, translação e rotação.

Vários métodos têm sido propostos para calcular a ordem da simetria rotacional e a orientação de uma dada forma, conforme pode ser visto em Leou & Tsai (1987). Nesse trabalho, Leou & Tsai descrevem um método para determinar a ordem da simetria rotacional de uma dada forma baseado apenas em operações simples em pixels e cálculo dos momentos de Segunda ordem.

13 – CONCLUSÕES

O presente apêndice procurou mostrar a flexibilidade e versatilidade do uso de momentos como descritores de características de formato, e que apresenta também comportamento invariante através de diversas possibilidades de normalizações, permitindo o seu uso em todos os tipos de imagem e aplicações, porém restrito à uma determinada aplicação onde as informações fornecidas sejam suficientes. O método dos Momentos proporciona uma técnica robusta para decompor uma imagem em um conjunto finito de características. O ponto forte dessa técnica é que ela é matematicamente simples, baseando-se na transformação linear direta. Sua principal limitação é a sensibilidade à ruído e também insensibilidade à detalhes específicos que não podem ser discriminados utilizando apenas tarefas de identificação de formas globais. É possível desenvolver algoritmos matemáticos que analisam silhuetas e produzem resultados compatíveis com a intuição humana. Técnicas baseadas no domínio espacial são bastante atrativas e populares.

⁹ Uma forma é denominada possuir simetria rotacional n -arestas (*n-fold*) se tornar-se igual a si mesma após ser rotacionada sobre seu centro de gravidade através de um múltiplo qualquer de $2\pi/n$. Similarmente, os eixos principais são indefinidos para muitas formas, tais como as formas rotacionalmente simétricas de n -arestas com $n \geq 3$ ou figuras simétricas à reflexão.

APÊNDICE 3

LÓGICA FUZZY

1 - INTRODUÇÃO

Dois dos principais aspectos da imperfeição da informação são a imprecisão e a incerteza. Estas duas características são intrinsecamente ligadas e opostas entre si: quanto mais se aumenta a incerteza mais se diminui a imprecisão e vice-versa. As teorias mais conhecidas para tratar da imprecisão e da incerteza são respectivamente a teoria dos conjuntos e a teoria de probabilidades. Estas teorias, embora muito úteis, nem sempre conseguem captar a riqueza da informação fornecida por seres humanos. A teoria dos conjuntos não é capaz de tratar o aspecto vago da informação e a teoria de probabilidades, na qual a probabilidade de um evento determina completamente a probabilidade do evento contrário, é mais adaptada para tratar de informação freqüentistas do que aquelas fornecidas por seres humanos (Correa, 1999).

A teoria dos conjuntos *fuzzy* foi desenvolvida a partir de 1965 por Lotfi Zadeh (Zadeh, 1965), para tratar do aspecto vago da informação. A teoria dos conjuntos, que será chamada neste trabalho de “clássica”, pode ser vista então como um caso particular desta teoria mais geral. A utilidade desse conjunto reside na sua habilidade de modelar dados incertos ou ambíguos encontrados freqüentemente na vida real (Pal & Mitra, 1992).

A partir de 1978, Lotfi Zadeh desenvolveu a teoria de possibilidades (Zadeh, 1978), que trata a incerteza da informação, podendo pois ser comparada com a teoria de probabilidades. Esta teoria, por ser menos restritiva, pode ser considerada mais adequada para o tratamento de informações fornecidas por seres humanos que a de probabilidades. Efetivamente, mesmo no discurso usual percebemos que a noção de possibilidade é menos restritiva que aquela de probabilidade: é mais fácil dizer que algum evento é possível do que provável.

A teoria dos conjuntos *fuzzy* e a teoria de possibilidades são intimamente ligadas. Por exemplo, o conjunto *fuzzy* que modela a informação “idade avançada” pode ser usada para modelar a distribuição de possibilidade da idade de uma dada pessoa, da qual só sabemos

que ela é idosa. O fato destas teorias serem ligadas é muito importante no sentido de que é possível se tratar tanto a imprecisão quanto a incerteza de um conjunto de informações em um único ambiente formal. De fato, a maior parte do tempo não é necessário fazer a distinção entre um conjunto *fuzzy* e uma distribuição de possibilidades. Estas teorias têm sido cada vez mais usadas em sistemas que utilizam informações fornecidas por seres humanos para automatizar procedimentos quaisquer, como por exemplo no controle de processos, no auxílio à decisão, etc. Estas teorias têm sido utilizadas em aplicações que vão do controle de eletrodomésticos ao controle de satélites, do mercado financeiro à medicina, e tendem a crescer cada vez mais, sobretudo em sistemas híbridos, que incorporam abordagens conexionistas e evolutivas (Correa, 1999).

A lógica *fuzzy* é uma das tecnologias atuais bem sucedidas para o desenvolvimento de sistemas para controlar processos sofisticados. Com sua utilização, requerimentos complexos podem ser Implementados em controladores simples, de fácil manutenção e baixo custo. O uso de sistemas construídos desta maneira, chamados de *controladores fuzzy*, é especialmente interessante quando o modelo matemático está sujeito a incertezas. Um controlador *fuzzy* é um sistema *fuzzy* à base de regras, composto de um conjunto de regras do tipo **Se** <premissa> **Então** <conclusão>, que definem ações de controle em função das diversas faixas de valores que as variáveis de estado do problema podem assumir (Klir & Folger, 1988). Estas faixas (usualmente mal definidas) de valores são modeladas por conjuntos *fuzzy* e denominados de termos lingüísticos. A maior dificuldade na criação de sistemas *fuzzy* em geral, e de controladores *fuzzy* em particular, encontra-se na definição dos termos lingüísticos e das regras.

Ao contrário dos controladores convencionais em que o algoritmo de controle é descrito analiticamente por equações algébricas ou diferenciais, através de um modelo matemático, em controle *fuzzy* utilizam-se regras lógicas no algoritmo de controle com a intenção de descrever numa rotina a experiência humana, intuição e heurística para controlar um processo (Zadeh, 1965). Os sistemas *fuzzy* são robustos, de grande adaptabilidade, incorporando conhecimentos que outros sistemas nem sempre conseguem acomodar. Também são versáteis quando o modelo físico é complexo e de difícil representação matemática.

Na próxima sessão serão abordados os principais conceitos e definições básicas relacionados com a teoria dos conjuntos *fuzzy*.

2 – CONCEITOS BÁSICOS DA TEORIA DOS CONJUNTOS FUZZY

A teoria dos conjuntos *fuzzy*, quando utilizada em um contexto lógico, como o de sistemas baseados em conhecimento, é conhecida como lógica *fuzzy*, lógica difusa ou lógica *fuzzy*. A principal diferença entre a proposição clássica e a *fuzzy* está na faixa de seus valores verdades (Klir & Yuan, 1995). Na teoria clássica um elemento pertence ou não a um determinado conjunto. Na teoria *fuzzy* o elemento pode pertencer, não pertencer ou estar parcialmente presente em um determinado conjunto. À cada elemento do conjunto *fuzzy* é dado um grau de pertinência. Esse valor de pertinência pertence a uma faixa de 0 (elemento não pertencente ao conjunto) até 1 (elemento totalmente pertencente ao conjunto). Uma função de pertinência é a relação entre os valores de um elemento e seu grau de pertinência em um conjunto. A figura A3.1 mostra uma função de pertinência trapezoidal, considerando a temperatura ambiente em torno de 23° C a 27° C. Nesta faixa a função de pertinência é igual a 1 indicando que a temperatura é ambiente, fora da faixa a temperatura será: quase ambiente (21° C a 29° C), não ambiente (0° C - 50° C), e assim por diante, dependendo do valor de sua pertinência.

Fig. A3.1 - Função de pertinência para temperatura ambiente.

O conjunto *fuzzy* permite representar conceitos vagos expressos na linguagem natural. A representação do conjunto *fuzzy* depende não somente do conceito, mas também do contexto no qual é utilizado (Klir & Yuan, 1995). Por exemplo, a temperatura alta no contexto de clima e no contexto do reator nuclear são representadas por conjuntos *fuzzy* diferentes.

Sendo x um elemento genérico de um conjunto X , um conjunto *fuzzy* de A em X (universo de discurso) é definido como um conjunto de pares ordenados:

$A = \{(x, \mu_A(x)) / x \in X\}$

onde $\mu_A(x)$ é chamada de função de pertinência de x em A . A função de pertinência mapeia cada elemento de X com um valor de pertinência contínuo entre 0 e 1 e indica o grau de compatibilidade entre x e o conceito expresso por A :

- $\mu_A(x) = 1$, indica que x é completamente compatível com A ;
- $\mu_A(x) = 0$, indica que x é completamente incompatível com A ;
- $0 < \mu_A(x) < 1$, indica que x é parcialmente compatível com A , com grau $\mu_A(x)$;

Um conjunto A da teoria dos conjuntos clássica pode ser visto como um conjunto *fuzzy* específico, denominado usualmente de “crisp” (definido), ou seja, a pertinência é do tipo “tudo ou nada”, “sim ou não”, e não gradual como para os conjuntos *fuzzy*. A diferença entre estes conceitos em relação à variável idade é ilustrada na figura A3.2a e na figura A3.2b, que descrevem respectivamente a representação do conceito “adolescente” através de um conjunto “crisp” e de um conjunto *fuzzy*. O conjunto “crisp” A não exprime completamente o conceito de “adolescente”, pois uma pessoa com 12 anos e 11 meses seria considerada completamente incompatível com este conceito. Na verdade, qualquer intervalo “crisp” que se tome para representar este conceito é arbitrário. Já o conjunto *fuzzy* B permite exprimir que qualquer pessoa com idade entre 13 e 17 anos é um adolescente, acima de 19 ou abaixo de 11 não é considerado um adolescente, e nos intervalos $[11,13]$ e $[17,19]$ é considerado tanto mais adolescente quanto mais próxima de 13 e respectivamente 17 anos é a sua idade.

Figura A3.2 – (a) Função característica do conjunto “crisp” adolescente. (b) Função trapezoidal característica do conjunto *fuzzy* adolescente (Correa, 1999).

Normalmente os valores de pertinência são definidos para variáveis lingüísticas (ex.: temperatura alta) através da quantificação da percepção humana (Ray & Ghoshal, 1997). A variável lingüística é caracterizada por uma quintupla $(x, T(x), U, G, M)$ na qual x é o nome da variável; $T(x)$ é o conjunto de termos de x , que é o conjunto de nomes dos valores lingüísticos de x no qual cada valor é um número *fuzzy* definido em U (universo de discurso); G é a regra sintática para gerar o nome dos valores de x e M é a regra semântica para associar cada valor com seu significado (Lee, 1990a). Por exemplo, se velocidade for interpretada como uma variável lingüística, então seus conjuntos de termos $T(\text{velocidade})$ poderiam ser:

$$T(\text{velocidade}) = \{ \text{lenta, moderada, rápida, muito lenta, mais ou menos rápida...} \}$$

Cada termo em $T(\text{velocidade})$ é caracterizado por um conjunto *fuzzy* no universo de discurso $U = [0, 150]$. O termo lento pode ser interpretado como a velocidade abaixo de 60 Km/h, o termo médio como a velocidade em torno de 80 Km/h e rápido como a velocidade acima de 120 Km/h. Esses termos podem ser caracterizados como o conjunto *fuzzy* no qual as funções de pertinências são mostradas na figura A3.3.

Figura A3.3 – Funções de pertinência para velocidade (França, 1999)

O uso de conjuntos *fuzzy* provê a base para manipulação sistemática de conceitos vagos e imprecisos, dentro de uma abordagem lingüística. A abordagem lingüística representa uma mistura entre as abordagens quantitativa e qualitativa, fazendo uso de palavras quanto a caracterização numérica não é apropriada, e usando número para tornar o significado das palavras mais precisos. A caracterização lingüística usa o conceito de valores que não são números, mas palavras de uma linguagem natural. Uma variável

lingüística é uma variável *fuzzy*, cujo valor, em um determinado momento, é um número *fuzzy* dentro do universo de discurso.

2.1 – Operações com Conjuntos Fuzzy

Similarmente às operações nos conjuntos “*crisp*”, existe a necessidade de proceder às operações de intersecção, união e negação, entre outras, nos conjuntos *fuzzy*.

Sejam A e B dois conjuntos *fuzzy* em U com suas funções de pertinências μ_A e μ_B , respectivamente. As operações de conjuntos como união, intersecção e complemento para conjuntos *fuzzy* são definidas através de suas funções de pertinências. Pode-se expressar a intersecção destes conjuntos, como um outro conjunto $E = A \cap B$. Da mesma forma, pode-se expressar a união como um conjunto $F = A \cup B$.

- União: A função de pertinência $\mu_{A \cup B}$ é definida ponto a ponto para todo $x \in U$ por:

$$\mu_{A \cup B}(x) = \max\{\mu_A(x), \mu_B(x)\} \quad \text{ou}$$

$$\mu_A(x) \cup \mu_B(x) = \mu_A(x) \vee \mu_B(x)$$

- Intersecção

A função de pertinência $\mu_{A \cap B}$ é definida ponto a ponto para todo $x \in U$ por:

$$\mu_{A \cap B}(x) = \min\{\mu_A(x), \mu_B(x)\} \quad \text{ou}$$

$$\mu_A(x) \cap \mu_B(x) = \mu_A(x) \wedge \mu_B(x)$$

- Complemento: A função de pertinência $\mu_{\neg A}$ é definida ponto a ponto para todo $x \in U$ por:

$$\mu_{\neg A}(x) = 1 - \mu_A(x)$$

A figura A3.4 mostra as operações de união e intersecção dos conjuntos *fuzzy* graficamente.

Figura A3.4 - Operação de união e intersecção nos conjuntos *fuzzy*. A área hachurada representa a intersecção e a linha mais forte nos contornos representa a união (França, 1999).

- Produto Cartesiano:

O produto cartesiano entre dois conjuntos *fuzzy* A e B é definido como:

$$\mu_A \times \mu_B = \min\{\mu_A(x), \mu_B(x)\}$$

- Relação *fuzzy*:

Sejam A e B dois conjuntos *fuzzy* não vazios. A relação *fuzzy* R é um subconjunto *fuzzy* de AxB. Se A = B, então R é uma relação binária *fuzzy* em A.

Um exemplo simples de uma relação binária *fuzzy*. Seja U = {1, 2, 3}. Se for definido o quanto um elemento do conjunto é igual ao outro, poderia-se ter:

$$\begin{aligned} R(1,1) &= R(2,2) = R(3,3) = 1 \\ R(1,2) &= R(2,1) = R(2,3) = R(3,2) = 0.7 \\ R(1,3) &= R(3,1) = 0.4 \end{aligned}$$

Em notação matricial, tem-se:

	1	2	3
1	1	0.7	0.4
2	0.7	1	0.7
3	0.4	0.7	1

Composição Sup-Star: Se R e S são relações *fuzzy* em $U \times V$ e $V \times W$, respectivamente, a composição R e S é uma relação *fuzzy* definida por:

$$RoS = \{(u,w), \sup [\mu_R(u,v) * \mu_S(v,w)] \} \qquad u \in U, v \in V e w \in W.$$

onde * pode ser qualquer operador da classe de norma triangular (geralmente min).

Exemplo: Sejam duas relações *fuzzy*:

R = x é maior que y =

	y1	y2	y3	y4
x1	0.8	0.1	0.1	0.7
x2	0	0.8	0	0
x3	0.9	1	0.8	0.8

S = y é fechado em z =

	z1	z2	z3
y1	0.4	0.9	0.3
y2	0	0.4	0
y3	0.9	0.5	0.8
y4	0.6	0.7	0.5

então a composição RoS é:

	z1	z2	z3
x1	0.6	0.8	0.5
x2	0	0.4	0
x3	0.7	0.9	0.7

Portanto, a composição RoS, nada mais é do que o produto cartesiano clássico de matrizes, sendo que no lugar da adição é utilizado o operador de união *fuzzy* e no lugar da multiplicação é utilizado o operador de intersecção *fuzzy*.

2.2 – Operadores de Intersecção (t - normas) e União (t- conormas)

Na teoria dos conjuntos *fuzzy*, a intersecção é implementada por uma família de operadores denominados de t-normas, e a união é implementada por uma família de operadores denominados t-conormas. É importante notar que as t-normas e t-conormas se

reduzem aos operadores clássicos de união e interseção quando os conjuntos são “crisp”. A tabela A3.1, adaptada de (Correa, 1999) e França (1999) mostra as t-normas e t-conormas duais básicas. A maior norma triangular é a intersecção e a menor é o produto drástico (Lee, 1990b). A maior t-conorma é a de Weber e a menor é máximo. Os operadores mais usualmente utilizados são os operadores de Zadeh e os probabilistas.

As operações associadas à norma triangular são definidas para todo $x,y \in [0,1]$. Essas normas são definidas como:

$$T:[0,1] \times [0,1] \rightarrow [0,1]$$

As t-conormas são definidas pelo seguinte mapeamento:

$$S:[0,1] \times [0,1] \rightarrow [0,1]$$

Tabela A3.1 - Normas e Conormas triangulares básicas

t-norma	t-conorma	Nome
<i>Interseção ou Mínimo</i> Min(a,b)	Max(a,b)	Zadeh
<i>Produto Algébrico ou Produto probabilístico</i> a . b	a + b – ab	Probabilista
<i>Produto Limitado</i> Max(a+b-1 , 0)	Min (a+b, 1)	Lukasiewicz
<i>Produto Drástico</i> a , se b=1 b, se a=1 0, senão	a , se b=0 b, se a=0 1, senão	Weber

A figura A3.5 ilustra alguns destes operadores em relação a dois conjuntos *fuzzy* A e B. Estes operadores satisfazem as leis de De Morgan em relação ao operador negação (Correa, 1999).

Figura A3.5 – Ilustração das principais t-normas e t-conormas (Correa, 1999)

2.3 - Implicação fuzzy

Os operadores de implicação são usados para modelar regras de inferências do tipo **SE** <premissa> **Então** <conclusão>. A implicação *fuzzy* (regras de inferência *fuzzy* ou condição de estado *fuzzy* ou simplesmente regra *fuzzy*) assume a seguinte forma:

$$SE\ x\ é\ A\ ENTÃO\ y\ é\ B\ ou\ A \rightarrow B$$

onde A e B são valores lingüísticos definidos por conjuntos *fuzzy* no universo de discurso X e Y, respectivamente. A implicação *fuzzy* se divide em duas partes, uma chamada de antecedente ou premissa (“x é A”) e outra denominada conseqüente ou conclusão (“y é B”). Exemplo de regras *fuzzy*:

SE pressão é alta ENTÃO volume é pequeno

SE tonalidade é escura E tamanho é grande ENTÃO atenção é grande

Os antecedentes de uma regra serão sempre proposições do tipo receptivo, ou seja, constataam uma determinada condição. Já os conseqüentes das regras podem ser de dois tipos. O primeiro deles é o tipo perceptivo, quando se atinge uma maior grau de abstração, como por exemplo na seguinte regra:

SE temperatura é *alta* E a pressão é *muito alta* ENTÃO a situação é *perigosa*.

Outro tipo de conseqüente é o tipo decisivo, ou seja, algum tipo de decisão tomada diante da constatação de uma determinada condição, como por exemplo na seguinte regra:

SE a situação é *perigosa*, ENTÃO abrir *bastante* a válvula de segurança.

A regra *fuzzy* pode ser definida como uma relação binária *fuzzy* R no produto do espaço X e Y, uma vez que descreve a relação entre duas variáveis x e y. Portanto, $A \rightarrow B$ deve ser definida ponto a ponto, e isso é feito através dos operadores de implicação *fuzzy* (tabela A3.2).

Tabela A3.2 - Operadores de implicação *fuzzy*

Operadores de implicação <i>fuzzy</i>	
Larsen	$x \rightarrow y = x \cdot y$
Lukasiewicz	$x \rightarrow y = \min\{1, 1 - x + y\}$
Mamdani	$x \rightarrow y = \min\{x, y\}$
Standart Strict Rescher-Gaines “Sharp”	$x \rightarrow y = 1, \text{ se } x \leq y$ 0 senão
Brower- Gödel	$x \rightarrow y = 1 \text{ se } x \leq y$ y senão
Gaines	$x \rightarrow y = 1 \text{ se } x \leq y$ $y/x \text{ c.c.}$
Kleene-Dienes	$x \rightarrow y = \max\{1-x, y\}$
Reichenbach “Estocástica”	$x \rightarrow y = 1 - x + xy$

A figura A3.6 mostra um exemplo:

Figura A3.6 - Exemplos de implicação fuzzy.

Se $x = 4 \Rightarrow$ pressão alta $\mu_A = 0.75$ e
se $y = 1 \Rightarrow$ volume pequeno $\mu_B = 1$

$A(4) \rightarrow B(1) = 0.75 \rightarrow 1$ através do operador Gödel
0.75 através do operador Larsen

2.4 - Raciocínio fuzzy

Raciocínio fuzzy é um procedimento de inferência usado para obter conclusão de um conjunto de regras IF-THEN de uma ou mais condições.

Há duas importantes regras de inferência fuzzy denominada *modus ponens* generalizada (GMP) e *modus tollens* generalizada (GMT) (Lee, 1990a), dada por:

fato: $x \text{ é } A'$
regra: $SE x \text{ é } A \text{ ENTÃO } y \text{ é } B$ (GMP)
consequência: $y \text{ é } B'$

fato: $y \text{ é } B'$
regra: $SE x \text{ é } A \text{ ENTÃO } y \text{ é } B$ (GMT)
consequência: $x \text{ é } A'$

A implicação fuzzy é baseada na regra composicional. O GMP se reduz ao *modus ponens* clássico quando $A'=A$ e $B'=B$, e está intimamente relacionado à inferência dirigida a dados, a qual é útil em lógica fuzzy de controle (Lee, 1990a). O GMT se reduz ao *modus*

tollens clássico quando $B' = \text{não } B$ e $A' = \text{não } A$, e está intimamente ligado à inferência dirigida a objetivos, o que é comumente usado em sistemas especialistas, especialmente na esfera de diagnósticos médicos (Lee, 1990a).

As regras *fuzzy* podem ser classificadas em três tipos dependendo de sua forma conseqüente (Lee, Kwak & Kwang, 1996) apud França (1999):

Tipo 1: regras *fuzzy* com a parte conseqüente constante

$R: \text{SE } x_1 \text{ é } A_1 \text{ e } \dots \text{ e } x_m \text{ é } A_m$
 $\text{ENTÃO } y \text{ é } C_i$

Tipo 2: regras *fuzzy* com a parte conseqüente sendo uma combinação linear

$R: \text{SE } x_1 \text{ é } A_1 \text{ e } \dots \text{ e } x_m \text{ é } A_m$
 $\text{ENTÃO } y \text{ é } g(x_1, \dots, x_m) = b_o + b_1 x_1 + \dots + b_m x_m$

Tipo 3: regras *fuzzy* com a parte conseqüente sendo um conjunto *fuzzy*

$R: \text{SE } x_1 \text{ é } A_1 \text{ e } \dots \text{ e } x_m \text{ é } A_m$
 $\text{ENTÃO } y \text{ é } B$

Nas regras acima, x_i e y denotam variáveis de entrada e saída respectivamente. Os termos lingüísticos são conjuntos *fuzzy* parametrizados, no qual a forma é determinada por poucos parâmetros (conjunto *fuzzy* triangular, trapezoidal, sigmoidal, entre outros). O termo c_i denota um valor constante, e $g(x_1 + \dots + x_m) = b_o + b_1 x_1 + \dots + b_m x_m$ é uma combinação linear das variáveis de entrada onde b_i 's são coeficientes constantes. Por fim, B indica um conjunto *fuzzy* parametrizado ou não parametrizado. Conjuntos *fuzzy* não parametrizados são definidos pela representação de seus valores e seu grau de pertinência, como exemplo temos:

$\text{velocidade baixa} = \{(0, 1.0), (80, 0.5), (120, 0.1)\}$

Podemos ter regras simples com antecedentes simples:

fato: $x \text{ é } A'$
regra: $\text{SE } x \text{ é } A \text{ ENTÃO } z \text{ é } C$
conseqüência: $z \text{ é } C'$

Ou podemos ter múltiplas regras com múltiplos antecedentes:

fato: $x \text{ é } A' \text{ e } y \text{ é } B'$

regra 1: SE x é $A1$ e y é $B1$ ENTÃO z é $C1$

regra 2: SE x é $A2$ e y é $B2$ ENTÃO z é $C2$

consequência: z é C'

Nas conexões AND (E) são utilizadas as normas triangulares e para realizar as ligações entre as diversas regras são utilizados as co-normas triangulares.

Ainda pode-se classificar as regras *fuzzy* quanto ao número de suas entradas e saídas. Nesse caso o sistema será referenciado como MIMO (múltiplas entradas e múltiplas saídas) e MISO (múltiplas entradas e saída simples) (Lee, 1990a).

Com o uso das variáveis lingüísticas e seus valores (termos) representados por conjuntos *fuzzy*, pode-se implementar o que se denomina uma proposição *fuzzy*. Uma proposição *fuzzy* é simplesmente a atribuição de um valor lingüístico a uma variável lingüística. Com o uso de preposições *fuzzys*, pode-se então montar a estrutura básica de representação do conhecimento em lógica *fuzzy* expressa por uma implicação *fuzzy* que, se uma determinada condição é satisfeita, uma conclusão pode ser tomada. Em lógica *fuzzy* a condição não precisa ser associada a uma certeza, podendo ser descrita por uma proposição *fuzzy* ou um conjunto de proposições *fuzzys*, que por sua própria definição contém elementos de incerteza (Politano, 1996).

2.5 - Modelos de Inferência *fuzzy*

O sistema de inferência *fuzzy* (FIS) é um sistema que utiliza a teoria dos conjuntos *fuzzy* para mapear entradas (características, no caso de classificador *fuzzy*) em saídas (classes, no caso do classificador) (Knapp, 1996). O procedimento de inferência é o núcleo de um sistema *fuzzy* e permite simular o mecanismo humano de tomada de decisão, empregando implicações *fuzzys* e regras de inferência da lógica *fuzzy*.

Segundo (Knapp, 1996) a saída de um modelo *fuzzy* se realiza através de seis passo:

- Determinação de um conjunto de regras *fuzzy* (criação das regras)
- “Fuzificação” das entradas usando as funções de pertinência
- Combinação das entradas “fuzificadas” em cada regra *fuzzy* para estabelecer a ativação da regra
- Cálculo da parte consequente da regra

- Combinação das partes conseqüente de todas as regras para obter a distribuição da saída
- “Desfuzificação” da saída.

2.5.1 – Base do Conhecimento

A base do conhecimento consiste de uma base de dados (funções) e uma base de regras, onde está contido o conhecimento sobre o domínio da aplicação e os objetivos a serem atingidos, de maneira a caracterizar a estratégia do processo e suas metas. Na base de dados ficam armazenadas as definições sobre discretizações e normalizações dos universos de discurso e as definições das funções de pertinência dos termos *fuzzy*.

2.5.1.1 – Determinação das variáveis de entrada (Conjuntos *fuzzy*)

As variáveis de entrada (conjuntos *fuzzy*) são n variáveis do tipo $i_1, i_2, i_3 \dots i_n$, que caracterizam, qualificam e alteram o processo (objeto do discurso) conforme ilustra a figura A3.7 . A identificação destas variáveis e como elas influenciam o processo é exaustivo para leigos mas é intuitivo para os especialistas que lidam diariamente com elas .

Fig. A3.7 - Variáveis de entrada (Borelli, 1999).

Para abstrair estes conhecimentos pode-se recorrer a manuais, experimentos ou questionar um especialista. Como exemplo da extração de conhecimentos por especialistas, recente trabalho apresentado por (Borelli, 1999) para caracterizar a beleza feminina, envolveu uma pesquisa com especialistas da área para a determinação dos conjuntos *fuzzy*. De acordo com a avaliação das respostas obtidas no questionário realizado para caracterizar a variável qualitativa beleza (objeto de discurso) determinou-se os seguintes conjuntos *fuzzy*: cintura, pernas, carisma, busto, peso e altura.

2.5.1.2 – Determinação das Funções de Pertinência (Base de Funções)

As funções de pertinências (*membership functions*) determinam o grau com que um determinado valor de entrada está presente dentro do conjunto *fuzzy* correspondente. Na “fuzificação”, as funções de pertinências mapeiam as entradas em valores pertencentes ao intervalo [0,1].

A base de funções contém as definições das funções de pertinência dos termos lingüísticos que cada variável pode assumir. Os conceitos definidos na base de funções são utilizados para caracterizar as regras *fuzzys* e a manipulação de dados *fuzzy*. Esses conceitos são subjetivos e definidos com base na experiência e no conhecimento de especialistas. Deve ser observado que a escolha correta das funções de pertinência de um conjunto de termos exerce um papel essencial para o sucesso da aplicação.(Politano, 1996).

Uma variável lingüística está associada a um conjunto de termos, onde cada termo do conjunto é definido dentro do mesmo universo de discurso. Uma partição *fuzzy* define quantos termos devem existir em um conjunto de termos, ou qual deve ser o número de conjuntos *fuzzy* primários. O número de conjuntos *fuzzy* é que determina a precisão ou o grau de pertinência. A figura A3.8 extraída de Politano, (1996) ilustra um exemplo típico, onde são mostradas dois tipos de partições *fuzzys* em um mesmo universo de discurso.

Figura A3.8 – Representação gráfica de dois tipos de partição *fuzzy*. (a) Partição com três termos: (Negativo, Zero, Positivo). (b) Partição com sete termos: Negativo Grande, Negativo Médio, Negativo pequeno, Zero, Positivo, Positivo Pequeno, Positivo Médio, Positivo Grande (Politano, 1996).

O número de elementos do conjunto de termos das variáveis lingüísticas de entrada determina o número máximo de regras que podem ser construídas. Por exemplo, no caso de um sistema *fuzzy* com duas entradas e uma saída, se o número de elementos são 3 para uma entrada e sete para a outra, o número máximo de regras é 3×7 . A definição do número de partições para variáveis lingüísticas não é determinística e não tem uma única solução, sendo necessária analisar diversas tentativas para se obter uma melhor partição.

Há dois métodos para se definir a função de pertinência de conjuntos *fuzzy*. Dependendo do universo de discurso ser discreto ou contínuo, o método usado é numérico ou por função respectivamente.

Os valores da função de pertinência de um conjunto *fuzzy* definido pelo método numérico, é representado por um vetor de números, cuja dimensão depende do universo de discretização. No método da definição da função, a função de pertinência dos conjuntos *fuzzy* pode ser feita através de funções matemáticas.

As funções mais utilizadas, descritas a seguir, estão ilustradas na figura A3.9:

Triangular: é especificada por três parâmetros $\{a,b,c\}$, a qual determina a coordenada x dos três cantos do triângulo.

$$Triangulo(x;a,b,c) = \max(0, \min[(x-a)/(b-a), (c-x)/(c-b)])$$

Trapezoidal: é especificada por quatro parâmetros $\{a,b,c,d\}$

$$Trap(x;a,b,c,d) = \max(0, \min[(x-a)/(b-a), 1, (d-x)/(d-c)])$$

Gaussiana: é especificada por dois parâmetros $\{\sigma,c\}$

$$Gaussiana(x; \sigma, c) = \exp\{-(x-c)/\sigma\}^2$$

Sino: é especificada por três parâmetros $\{a,b,c\}$

$$Sino(x;a,b,c) = 1/(1 + [(x-c)/a]^{2b})$$

Sigmoidal: é especificada por dois parâmetros $\{a,c\}$

$$Sig(x;a,c) = 1/(1 + \exp[-a(x-c)])$$

Figura A3.9 - Funções de pertinência utilizadas para "fuzificação". (a) função triangular; (b) função trapezoidal; (c) função Gaussiana; (d) função seno e (e) função sigmoidal (França, 1999).

As funções de pertinência são definidas, para cada aplicação, de acordo com os termos necessário de cada conjunto *fuzzy*. No trabalho apresentado por (Borelli, 1999), o conjunto *fuzzy* Peso, por exemplo foi dividido em três termos lingüísticos: Pouco Peso, Médio Peso e Muito Peso. Por isso, são necessárias três funções de pertinência, uma para cada termo.

Para descrever o tipo de função de pertinência que mais adequada (Triângulo, Trapézio, Gaussiana, etc..) deve-se recorrer novamente ao conhecimento de especialistas da área ou à pesquisa minuciosa. As funções de pertinências definidas funcionalmente podem ser facilmente redefinidas quando ocorrem mudanças no universo de discurso ou durante o processo de melhoria da especialização das funções.

Borelli, (1999) elaborou também, um outro questionário com o objetivo de determinar a função mais adequada para representação das funções de pertinências de cada

conjunto *fuzzy*. Para cada uma das funções de pertinência fez-se um grupo de perguntas, para saber a frequência ou peso das respostas. A tabela A3.3 a seguir mostra as respostas obtidas dos especialistas para caracterizar o termo “pouca cintura”.

Tabela A3.3 – Dados obtidos para o termo “pouca cintura”

Pouca Cintura (cm)	Porcentagem de Votos
40	27 %
45	26 %
50	18 %
60	14 %
65	8 %
70	4 %
75	1 %

Os dados acima resultaram no gráfico mostrado na figura A3.10.

Figura A3.10 – Gráfico resultante da avaliação do questionário para o termo “Pouca Cintura” (Borelli, 1999).

Pelo gráfico mostrado na figura A3.10, a função que melhor se ajusta ao termo “Pouca Cintura” é a função trapézio mostrada na figura A3.11 a seguir:

Figura A3.11 - Função de pertinência para o termo Pouca Cintura (Borelli, 1999).

Para os outros termos lingüísticos do conjunto *fuzzy* Cintura (tabela 6), foram realizados os questionários resultando no gráfico da figura A3.12.

Figura A3.12 - Representação gráfica das funções de pertinências do conjunto *fuzzy* Cintura.

2.5.1.3 – Determinação da Base de Regras e do(s) conjunto(s) *fuzzy* de saída

Para a criação da base de regras *fuzzy* primeiramente é necessário identificar as variáveis de entrada e seus respectivos valores, as quais caracterizam o estado do sistema (variáveis de estado) e as normalizam em um universo de discurso padronizados. Estes valores são então “fuzificados”, com a transformação da entrada “*crisp*” em conjuntos *fuzzy*, para que possam se tornar instâncias de variáveis lingüísticas.

As regras *fuzzy* são mais convenientemente formuladas em termos lingüísticos do que numéricos. Sob essa ótica, as regras *fuzzy* provêm uma estrutura natural para a caracterização do comportamento humano no processo de tomada de decisão. Muitos

especialistas descobriram que as regras *fuzzy* são uma forma conveniente de expressar seu domínio de conhecimento.

A escolha correta das variáveis do antecedente e das variáveis do conseqüente é essencial para a caracterização da operação de um sistema *fuzzy*. Além disso, a seleção das variáveis lingüísticas tem um efeito significativo no desempenho do sistema. Conforme visto anteriormente, a experiência e conhecimento do especialista exercem um papel importante durante o estágio da seleção das variáveis.

As regras *fuzzy*, juntamente com os dados de entrada, são processadas pelo procedimento de inferência, o qual infere as ações de classificação ou controle, de acordo com o estado do sistema, aplicando o operador de implicação. São escritas na seguinte forma:

SE (entrada 1)... E/OU (entrada 2)... ENTÃO (saída)

Em um dado sistema *fuzzy* é importante haver tantas regras quantos forem necessárias para mapear totalmente as combinações dos termos das variáveis (base completa), garantindo que exista sempre pelo menos uma regra a ser disparada para qualquer entrada. Também são essenciais a consistência, onde procura-se evitar a possibilidade de contradições, e a interação entre as regras gerenciada pela função de implicação.

As regras de inferência representam o conhecimento dos especialistas, resultados de pesquisa ou sistema de avaliação. O número de regras é dado pelo produto escalar das funções de pertinência de cada conjunto *fuzzy*. No trabalho de (Borelli, 1999) os conjuntos fuzzy Peso (pouco, médio, muito), Cintura (pouco, médio, muito), Pernas (pouco, muito), Carisma (pouco, médio, muito) e Altura (pouco, médio, muito) resultaram em 324 regras resultantes da interação de suas funções de pertinência.

Para a elaboração das regras há duas abordagens empíricas que dependem basicamente da origem do conhecimento. A abordagem mais comum envolve uma introspecção e uma verbalização de modo a formalizar o conhecimento do especialista. Uma outra abordagem compreende a elaboração de uma série de pesquisas, questionários, livros, relatórios observação do fenômeno, com o objetivo de averiguar o relacionamento das variáveis lingüísticas de entrada com a variável de saída.

Como a obtenção das regras *fuzzys* está baseada na experiência ou no conhecimento de especialistas, as regras podem estar sujeitas a diferentes critérios de desempenho. Na prática, é importante verificar a consistência das regras de modo a minimizar a possibilidade de contradições.

Como um outro exemplo imagine que deseja-se saber a proporcionalidade entre peso e altura das candidatas para uma pré-classificação do concurso de beleza. Para isto, consideremos dois conjuntos *fuzzy*: Peso e Altura e as funções de pertinência de entrada: Peso (Pouco, Médio, Muito) e Altura (Pouco, Médio, Muito). Na saída consideremos o conjunto *fuzzy*: Proporcionalidade, com as seguintes funções de pertinência: Pouco e Muito.

Pesquisar, significa achar resposta para perguntas do tipo:

Se o Peso é Pouco e a Altura é Pouco, o que ocorre com a Proporcionalidade?

Se o Peso é Pouco e a Altura é Médio, o que ocorre com a Proporcionalidade?

Se o Peso é Pouco e a Altura é Muito, O que ocorre com a Proporcionalidade?

Se o Peso é Médio e a Altura é Pouco, o que ocorre com a Proporcionalidade?

Se o Peso é Médio e a Altura é Médio, o que ocorre com a Proporcionalidade?

Se o Peso é Médio e a Altura é Muito, O que ocorre com a Proporcionalidade?

Se o Peso é Muito e a Altura é Pouco, o que ocorre com a Proporcionalidade?

Se o Peso é Muito e a Altura é Médio, o que ocorre com a Proporcionalidade?

Se o Peso é Muito e a Altura é Muito, o que ocorre com a Proporcionalidade?

Para responder a estas nove perguntas deve-se recorrer a pesquisa: catálogos, livros, ou ainda extrair o conhecimento e experiência de especialistas através de questionários. Com as respostas pode-se montar as regras de inferência *fuzzy*. Para este exemplo proposto as regras são dadas pela tabela A3.4.

Tabela A3.4 - Regras de inferência para a variável de saída “ Proporcionalidade” (Borelli, 1999)

SE	ENTÃO
Se (peso é pouco) E altura é (pouco)	Então (proporcionalidade é pouco)
Se (peso é pouco) E altura é (médio)	Então (proporcionalidade é muito)
Se (peso é pouco) E altura é (muito)	Então (proporcionalidade é muito)
Se (peso é médio) E altura é (pouco)	Então (proporcionalidade é pouco)
Se (peso é médio) E altura é (médio)	Então (proporcionalidade é pouco)
Se (peso é médio) E altura é (muito)	Então (proporcionalidade é muito)
Se (peso é muito) E altura é (pouco)	Então (proporcionalidade é pouco)
Se (peso é muito) E altura é (médio)	Então (proporcionalidade é pouco)
Se (peso é muito) E altura é (muito)	Então (proporcionalidade é médio)

A determinação dos conjuntos *fuzzy* de saída corresponde a identificação do objeto do discurso, ou seja, a resposta desejada.

Para avaliar a aptidão das candidatas ao concurso de beleza, (Borelli, 1999) utilizou a variável qualitativa “Beleza”. Para a representar o conjunto *fuzzy* Beleza, foram escolhidas os seguintes termos lingüísticos baseados nas informações dos especialistas: Pouquíssimo, Pouco, Médio, Bom, Muito, MUITÍSSIMO. Todas as respostas das 324 perguntas foram analisadas. Os resultados foram registrados para a elaboração das regras de inferência. A função triangular foi a que melhor se ajustou para representação das funções de pertinência do conjunto de saída *fuzzy* Beleza.

Quanto menor a base do triângulo maior é a precisão da resposta. A figura A3.13 mostra as funções de pertinência para o conjunto *fuzzy* Beleza obtidas por Borelli (1999).

Quando o número de regras é muito grande, pode ser usada uma *look_up table* que guarde o endereço das regras desejadas, ou seja, as que melhor se encaixam ao problema.

Figura A3.13 – Representação das função de pertinência de “Beleza” (Borelli, 1999).

2.1.5.4 - "Fuzificação"

A "fuzificação" é a transformação dos valores observados em valores *fuzzy* através do uso de funções de pertinências. A “fuzificação” toma os valores das variáveis de entrada, faz um escalonamento para condicionar os valores de entrada para seus valores correspondentes, dentro do universo de discurso e “fuzifica” os valores, transformando simples números em conjuntos *fuzzy*, de modo que possam se tornar instâncias de variáveis lingüísticas. A função da “fuzificação” está relacionada aos aspectos vagos e imprecisos da linguagem natural e consiste em avaliar subjetivamente um dado de entrada.

Na avaliação da beleza de uma candidata (Borelli, 1999), o valor de seu peso terá, no conjunto *fuzzy* Peso, três valores relacionado aos termos lingüísticos: “Peso Pouco”, “Peso Médio” e “Peso Muito”. Este valor depende da função de pertinência e é chamado valor de pertinência.

Como exemplo, na figura A3.14, uma candidata de 60 kg tem 70,833% de “Peso Médio” e 29,166% de “Peso Pouco” e 0% de “Peso Muito”. Em outras palavras, a candidata de 60 kg terá um valor de pertinência de 0.70833 para o termo lingüístico “Peso Médio”, um valor 0.29166 para o termo “Peso Pouco” e 0.0 para o termo “Peso Muito”.

Figura A3.14 - "Fuzificação" do valor de entrada de 60 kg dentro da variável *fuzzy* *Peso*.

Nas regras *fuzzy* são utilizados os conceitos AND e OR e para implementá-las são utilizados as t-normas e as t-conormas, respectivamente.

A parte conseqüente de cada regra é conseguida através do uso de algum operador de implicação. Os mais utilizados são a de Mandami e Larsen.

As saídas de todas as regras devem ser *combinadas* para obter a distribuição de saída *fuzzy*. Geralmente é utilizado o operador de união.

2.1.5.6 - "Desfuzificação"

Em muitos casos é desejável que a saída do sistema não seja *fuzzy* e sim uma saída definida. A etapa de "desfuzificação" mapeia a saída *fuzzy* em um valor definido (*crisp*).

O objetivo da "desfuzificação" é obter uma conclusão precisa que represente, da melhor maneira possível, as conclusões *fuzzys* restantes do procedimento de inferência. Algumas estratégias de "desfuzificação" são:

Método do máximo critério (MAX): O máximo critério acha o ponto no qual a distribuição de saída possui o seu primeiro valor de máximo (figura A3.15).

Método da média dos máximos (MOM): O MOM acha o valor médio onde a saída possui seus máximos valores. Isto pode ser computado como (figura A3.15):

$$z_o = \sum (w_j / l)$$

onde w_j são os valores no qual a função de pertinência possui seu máximos valores e l é o número de valores máximos encontrados.

Método de centro de massa (COA): É largamente usado. Essa estratégia procura pelo centro de gravidade da distribuição de possibilidades de uma conclusão (saída *fuzzy*). (Lee, 1990b).

Figura A3.15- Estratégias de "desfuzificação". Método do máximo critério (MAX). Método da média dos máximos(MOM). Método do centro de massa (COA).

2. 6 - Alguns modelos *fuzzy*

Alguns modelos *fuzzy* são largamente utilizadas em várias aplicações. Geralmente as diferenças entre os sistemas está na sua parte conseqüente, no seu processo de agregação das várias regras e na “desfuzificação” do sistema.

Figura A3.16 - Raciocínio *fuzzy* de Mandami (França, 1999).

2.6.2 - Modelo *fuzzy* de Larsen

É uma variação do modelo de Mandami frequentemente encontrado na literatura. Ao invés de se utilizar o operador de Mandami nas implicações *fuzzy* das regras, é utilizado o operador de Larsen (produto). Portanto o modelo *fuzzy* de Mandami pode ter variações se forem feitas diferentes escolhas para as conectivas AND (t-normas) e OR (t-conormas). A figura A3.17 mostra o raciocínio *fuzzy* do modelo de Larsen.

Figura A3.17 - Modelo *fuzzy* de Larsen (França, 1999).

Nesses dois modelos se for necessário uma saída definida, utiliza-se uma etapa de “desfuzificação” (método do máximo critério, método da média máxima ou método do centro de massa).

2.6.3 - Modelo de Sugeno

Nos modelos de interpolação, cada conclusão é dada através de uma função estritamente monotônica, usualmente diferente para cada regra. No modelo de Takagi-Sugeno, a função é uma combinação linear das entradas tendo como parâmetros um conjunto de constantes.

Sugeno e Takagi (Sugeno & Takagi, 1985) usaram a seguinte arquitetura:

Regra 1: $SE\ x\ é\ A1\ e\ y\ é\ B1\ ENTÃO\ z1 = a_1x + b_1y + r_1$

Regra 2: SE x é A_2 e y é B_2 ENTÃO $z_2 = a_2x + b_2y + r_2$

fato: $x \acute{e} x o' e y \acute{e} y o'$

conseqüência.: z é zo

A ativação da regra é dada por:

$$\alpha_1 = A_1(x_o) \wedge B_1(y_o) \quad \alpha_2 = A_2(x_o) \wedge B_2(y_o)$$

A saída de cada regra é conseguida através de uma relação polinomial:

$$z_1 = a_1 x_o + b_1 y_o + r_1 \quad z_2 = a_2 x_o + b_2 y_o + r_2$$

Na agregação das regras, feita através do centro de massa, é conseguido uma saída definida (figura A3.18).

Figura A3.18 - Modelo *fuzzy* de Sugeno (França, 1999).

2.6.4 - Modelo *fuzzy* de Tsukamoto

No esquema de de Tsukamoto, a função é geralmente não linear, tendo como domínio os possíveis graus de compatibilidade entre cada premissa e as entradas. Neste esquema e também no modelo de Takagi-Sugeno, obtém-se para cada regra, um único valor

para a variável de controle, Finalmente, uma ação de controle global é obtida fazendo-se uma média ponderada dos valores individuais obtidos, onde cada peso é o próprio graus de compatibilidade entre a premissa da regra e as entradas, normalizado.

Neste modelo a todas as funções de pertinência são monotônicas. O centro de massa é utilizado para obter a saída, que no caso já é definida. A figura A3.19 mostra o raciocínio deste modelo.

Figura 7.19 - Modelo *fuzzy* de Tsukamoto (França, 1999).

Os modelos *fuzzy* de Mandami e Larsen exigem técnicas de “desfuzificação” caso uma saída definida seja desejada, enquanto que os modelos de Tsukamoto e Sugeno já fornecem a saída definida. Segundo Jang (Jang & Gulley, 1995) as vantagens dos modelos de Sugeno e Tsukamoto são: eficiência computacional, trabalha bem com técnicas lineares (controle PID), garante a continuidade da superfície de saída e é adequado a análise matemática.

3 – CONCLUSÃO

Existe uma incerteza proveniente dos processos humanos do pensamento, raciocínio, aprendizagem e percepção. A aprendizagem e percepção de um ambiente físico através dos nossos sensores naturais (olhos, nariz, ouvidos, etc), a percepção de dor e de outros eventos biológicos similares pelo sistema nervoso e redes neuronais requer atenção especial. O fenômeno de percepção associado com esses processos estão repletos de incertezas e não podem ser caracterizados através da teoria estatística convencional (Gupta, 1991).

O sistema lógico clássico é poderoso para resolver problemas especificados em termos de dois valores (binário). O sistema lógico *fuzzy*, como extensão dos sistemas clássicos consegue manipular problemas em que há vários estados de decisão, incluindo o binário. Além disso, expressamos nossa percepção do ambiente físico lingüisticamente e o sistema *fuzzy* trata com conceitos lingüísticos (mais, menos, maior,...) de uma forma parecida com a do ser humano.

As principais vantagens do uso da lógica *fuzzy* são (Custódio et al., 1994):

- permite a representação do conhecimento e inferência dos especialistas na forma de regras *fuzzys*, que são muito próximas da linguagem natural.
- Simplificação do desenvolvimento, manutenção e experimentação, uma vez que é muito mais fácil mudar regras do que programas.
- Representação de conceitos vagos, isto é, incorpora e pondera a incerteza do sistema.

O procedimento de raciocínio *fuzzy* é paralelo, ou seja, a combinação da parte antecedente da regra com a entrada pode ser computada de forma paralela. Isso faz com que o procedimento *fuzzy* seja apropriado para ser implementado em processadores paralelo

Para a execução deste trabalho, optou-se pelo modelo de Mandami. As vantagens dos modelos de Mandami e Larsen são, segundo Jang (Jang & Gulley, 1995): mais intuitivo, larga divulgação e é mais adequado às entradas humanas. O modelo de Mandami trunca as funções de pertinência nas partes conseqüentes das regras através do operador MIN. O modelo de Larsen diminui a escala das funções de pertinência das saídas das regras usando o operador do produto. O valor de saída das regras entre os dois modelos podem ser iguais ou diferentes dependendo do método de “desfuzificação” utilizado.

APÊNDICE 4

TABELAS DE REGRAS DEFINIDAS PARA AS TAREFAS DO
SUBSISTEMA *TOP-DOWN*

Tabela A4.1 – Tabela de Regras para a tarefa de determinar os objetos “escuros”

REGRA	CINZA	TAMANHO	ORIENTAÇÃO	FORMA	DISTÂNCIA	SALIÊNCIA TOPDOWN
1	Pouca	Pouca	Pouca	Cir	Pouca	Muita
2	Pouca	Pouca	Pouca	Cir	Muita	Muita
3	Pouca	Pouca	Pouca	Cci	Pouca	Muita
4	Pouca	Pouca	Pouca	Cci	Muita	Muita
5	Pouca	Pouca	Pouca	Ali	Pouca	Muita
6	Pouca	Pouca	Pouca	Ali	Muita	Muita
7	Pouca	Pouca	Pouca	Ccc	Pouca	Muita
8	Pouca	Pouca	Pouca	Ccc	Muita	Muita
9	Pouca	Pouca	Pouca	Chb	Pouca	Muita
10	Pouca	Pouca	Pouca	Chb	Muita	Muita
11	Pouca	Pouca	Pouca	Chi	Pouca	Muita
12	Pouca	Pouca	Pouca	Chi	Muita	Muita
13	Pouca	Pouca	Pouca	Hll	Pouca	Muita
14	Pouca	Pouca	Pouca	Hll	Muita	Muita
15	Pouca	Pouca	Muita	Cir	Pouca	Muita
16	Pouca	Pouca	Muita	Cir	Muita	Muita
17	Pouca	Pouca	Muita	Cci	Pouca	Muita
18	Pouca	Pouca	Muita	Cci	Muita	Muita
19	Pouca	Pouca	Muita	Ali	Pouca	Muita
20	Pouca	Pouca	Muita	Ali	Muita	Muita
21	Pouca	Pouca	Muita	Ccc	Pouca	Muita
22	Pouca	Pouca	Muita	Ccc	Muita	Muita
23	Pouca	Pouca	Muita	Chb	Pouca	Muita
24	Pouca	Pouca	Muita	Chb	Muita	Muita
25	Pouca	Pouca	Muita	Chi	Pouca	Muita
26	Pouca	Pouca	Muita	Chi	Muita	Muita
27	Pouca	Pouca	Muita	Hll	Pouca	Muita
28	Pouca	Pouca	Muita	Hll	Muita	Muita
29	Pouca	Média	Pouca	Cir	Pouca	Muita
30	Pouca	Media	Pouca	Cir	Muita	Muita
31	Pouca	Media	Pouca	Cci	Pouca	Muita
32	Pouca	Media	Pouca	Cci	Muita	Muita
33	Pouca	Media	Pouca	Ali	Pouca	Muita
34	Pouca	Media	Pouca	Ali	Muita	Muita
35	Pouca	Media	Pouca	Ccc	Pouca	Muita
36	Pouca	Media	Pouca	Ccc	Muita	Muita
37	Pouca	Media	Pouca	Chb	Pouca	Muita
38	Pouca	Media	Pouca	Chb	Muita	Muita
39	Pouca	Media	Pouca	Chi	Pouca	Muita

REGRA	CINZA	TAMANHO	ORIENTAÇÃO	FORMA	DISTÂNCIA	SALIÊNCIA TOPDOWN
40	Pouca	Media	Pouca	Chi	Muita	Muita
41	Pouca	Media	Pouca	Hll	Pouca	Muita
42	Pouca	Media	Pouca	Hll	Muita	Muita
43	Pouca	Media	Muita	Cir	Pouca	Muita
44	Pouca	Media	Muita	Cir	Muita	Muita
45	Pouca	Media	Muita	Cci	Pouca	Muita
46	Pouca	Media	Muita	Cci	Muita	Muita
47	Pouca	Media	Muita	Ali	Pouca	Muita
48	Pouca	Media	Muita	Ali	Muita	Muita
49	Pouca	Media	Muita	Ccc	Pouca	Muita
50	Pouca	Media	Muita	Ccc	Muita	Muita
51	Pouca	Media	Muita	Chb	Pouca	Muita
52	Pouca	Media	Muita	Chb	Muita	Muita
53	Pouca	Media	Muita	Chi	Pouca	Muita
54	Pouca	Media	Muita	Chi	Muita	Muita
55	Pouca	Media	Muita	Hll	Pouca	Muita
56	Pouca	Media	Muita	Hll	Muita	Muita
57	Pouca	Muita	Pouca	Cir	Pouca	Muita
58	Pouca	Muita	Pouca	Cir	Muita	Muita
59	Pouca	Muita	Pouca	Cci	Pouca	Muita
60	Pouca	Muita	Pouca	Cci	Muita	Muita
61	Pouca	Muita	Pouca	Ali	Pouca	Muita
62	Pouca	Muita	Pouca	Ali	Muita	Muita
63	Pouca	Muita	Pouca	Ccc	Pouca	Muita
64	Pouca	Muita	Pouca	Ccc	Muita	Muita
65	Pouca	Muita	Pouca	Chb	Pouca	Muita
66	Pouca	Muita	Pouca	Chb	Muita	Muita
67	Pouca	Muita	Pouca	Chi	Pouca	Muita
68	Pouca	Muita	Pouca	Chi	Muita	Muita
69	Pouca	Muita	Pouca	Hll	Pouca	Muita
70	Pouca	Muita	Pouca	Hll	Muita	Muita
71	Pouca	Muita	Muita	Cir	Pouca	Muita
72	Pouca	Muita	Muita	Cir	Muita	Muita
73	Pouca	Muita	Muita	Cci	Pouca	Muita
74	Pouca	Muita	Muita	Cci	Muita	Muita
75	Pouca	Muita	Muita	Ali	Pouca	Muita
76	Pouca	Muita	Muita	Ali	Muita	Muita
77	Pouca	Muita	Muita	Ccc	Pouca	Muita
78	Pouca	Muita	Muita	Ccc	Muita	Muita
79	Pouca	Muita	Muita	Chb	Pouca	Muita
80	Pouca	Muita	Muita	Chb	Muita	Muita
81	Pouca	Muita	Muita	Chi	Pouca	Muita
82	Pouca	Muita	Muita	Chi	Muita	Muita
83	Pouca	Muita	Muita	Hll	Pouca	Muita
84	Pouca	Muita	Muita	Hll	Muita	Muita
85	Média	Pouca	Pouca	Cir	Pouca	media
86	Media	Pouca	Pouca	Cir	Muita	media
87	Media	Pouca	Pouca	Cci	Pouca	media
88	Media	Pouca	Pouca	Cci	Muita	media
89	Media	Pouca	Pouca	Ali	Pouca	media

REGRA	CINZA	TAMANHO	ORIENTAÇÃO	FORMA	DISTÂNCIA	SALIÊNCIA TOPDOWN
90	Media	Pouca	Pouca	Ali	Muita	media
91	Media	Pouca	Pouca	Ccc	Pouca	media
92	Media	Pouca	Pouca	Ccc	Muita	media
93	Media	Pouca	Pouca	Chb	Pouca	media
94	Media	Pouca	Pouca	Chb	Muita	media
95	Media	Pouca	Pouca	Chi	Pouca	media
96	Media	Pouca	Pouca	Chi	Muita	media
97	Media	Pouca	Pouca	Hll	Pouca	media
98	Media	Pouca	Pouca	Hll	Muita	media
99	Media	Pouca	Muita	Cir	Pouca	media
100	Media	Pouca	Muita	Cir	Muita	media
101	Media	Pouca	Muita	Cci	Pouca	media
102	Media	Pouca	Muita	Cci	Muita	media
103	Media	Pouca	Muita	Ali	Pouca	media
104	Media	Pouca	Muita	Ali	Muita	media
105	Media	Pouca	Muita	Ccc	Pouca	media
106	Media	Pouca	Muita	Ccc	Muita	media
107	Media	Pouca	Muita	Chb	Pouca	media
108	Media	Pouca	Muita	Chb	Muita	media
109	Media	Pouca	Muita	Chi	Pouca	media
110	Media	Pouca	Muita	Chi	Muita	media
111	Media	Pouca	Muita	Hll	Pouca	media
112	Media	Pouca	Muita	Hll	Muita	media
113	Média	Média	Pouca	Cir	Pouca	media
114	Media	Media	Pouca	Cir	Muita	media
115	Media	Media	Pouca	Cci	Pouca	media
116	Media	Media	Pouca	Cci	Muita	media
117	Media	Media	Pouca	Ali	Pouca	media
118	Media	Media	Pouca	Ali	Muita	media
119	Media	Media	Pouca	Ccc	Pouca	media
120	Media	Media	Pouca	Ccc	Muita	media
121	Media	Media	Pouca	Chb	Pouca	media
122	Media	Media	Pouca	Chb	Muita	media
123	Media	Media	Pouca	Chi	Pouca	media
124	Media	Media	Pouca	Chi	Muita	media
124	Media	Media	Pouca	Hll	Pouca	media
126	Media	Media	Pouca	Hll	Muita	media
127	Media	Media	Muita	Cir	Pouca	media
128	Media	Media	Muita	Cir	Muita	media
129	Media	Media	Muita	Cci	Pouca	media
130	Media	Media	Muita	Cci	Muita	media
131	Media	Media	Muita	Ali	Pouca	media
132	Media	Media	Muita	Ali	Muita	media
133	Media	Media	Muita	Ccc	Pouca	media
134	Media	Media	Muita	Ccc	Muita	media
135	Media	Media	Muita	Chb	Pouca	media
136	Media	Media	Muita	Chb	Muita	media
137	Media	Media	Muita	Chi	Pouca	media
138	Media	Media	Muita	Chi	Muita	media
139	Media	Media	Muita	Hll	Pouca	media

REGRA	CINZA	TAMANHO	ORIENTAÇÃO	FORMA	DISTÂNCIA	SALIÊNCIA TOPDOWN
140	Media	Media	Muita	Hll	Muita	media
141	Media	Muita	Pouca	Cir	Pouca	media
142	Media	Muita	Pouca	Cir	Muita	media
143	Media	Muita	Pouca	Cci	Pouca	media
144	Media	Muita	Pouca	Cci	Muita	media
145	Media	Muita	Pouca	Ali	Pouca	media
146	Media	Muita	Pouca	Ali	Muita	media
147	Media	Muita	Pouca	Ccc	Pouca	media
148	Media	Muita	Pouca	Ccc	Muita	media
149	Media	Muita	Pouca	Chb	Pouca	media
150	Media	Muita	Pouca	Chb	Muita	media
151	Media	Muita	Pouca	Chi	Pouca	media
152	Media	Muita	Pouca	Chi	Muita	media
153	Media	Muita	Pouca	Hll	Pouca	media
154	Media	Muita	Pouca	Hll	Muita	media
155	Media	Muita	Muita	Cir	Pouca	media
156	Media	Muita	Muita	Cir	Muita	media
157	Media	Muita	Muita	Cci	Pouca	media
158	Media	Muita	Muita	Cci	Muita	media
159	Media	Muita	Muita	Ali	Pouca	media
160	Media	Muita	Muita	Ali	Muita	media
161	Media	Muita	Muita	Ccc	Pouca	media
162	Media	Muita	Muita	Ccc	Muita	media
163	Media	Muita	Muita	Chb	Pouca	media
164	Media	Muita	Muita	Chb	Muita	media
165	Media	Muita	Muita	Chi	Pouca	media
166	Media	Muita	Muita	Chi	Muita	media
167	Media	Muita	Muita	Hll	Pouca	media
168	Media	Muita	Muita	Hll	Muita	media
169	Muita	Pouca	Pouca	Cir	Pouca	pouca
170	Muita	Pouca	Pouca	Cir	Muita	pouca
171	Muita	Pouca	Pouca	Cci	Pouca	pouca
172	Muita	Pouca	Pouca	Cci	Muita	pouca
173	Muita	Pouca	Pouca	Ali	Pouca	pouca
174	Muita	Pouca	Pouca	Ali	Muita	pouca
175	Muita	Pouca	Pouca	Ccc	Pouca	pouca
176	Muita	Pouca	Pouca	Ccc	Muita	pouca
177	Muita	Pouca	Pouca	Chb	Pouca	pouca
178	Muita	Pouca	Pouca	Chb	Muita	pouca
179	Muita	Pouca	Pouca	Chi	Pouca	pouca
180	Muita	Pouca	Pouca	Chi	Muita	pouca
181	Muita	Pouca	Pouca	Hll	Pouca	pouca
182	Muita	Pouca	Pouca	Hll	Muita	pouca
183	Muita	Pouca	Muita	Cir	Pouca	pouca
184	Muita	Pouca	Muita	Cir	Muita	pouca
185	Muita	Pouca	Muita	Cci	Pouca	pouca
186	Muita	Pouca	Muita	Cci	Muita	pouca
187	Muita	Pouca	Muita	Ali	Pouca	pouca
188	Muita	Pouca	Muita	Ali	Muita	pouca
189	Muita	Pouca	Muita	Ccc	Pouca	pouca

REGRA	CINZA	TAMANHO	ORIENTAÇÃO	FORMA	DISTÂNCIA	SALIÊNCIA TOPDOWN
190	Muita	Pouca	Muita	Ccc	Muita	pouca
191	Muita	Pouca	Muita	Chb	Pouca	pouca
192	Muita	Pouca	Muita	Chb	Muita	pouca
193	Muita	Pouca	Muita	Chi	Pouca	pouca
194	Muita	Pouca	Muita	Chi	Muita	pouca
195	Muita	Pouca	Muita	Hll	Pouca	pouca
196	Muita	Pouca	Muita	Hll	Muita	pouca
197	Muita	Média	Pouca	Cir	Pouca	pouca
198	Muita	Media	Pouca	Cir	Muita	pouca
199	Muita	Media	Pouca	Cci	Pouca	pouca
200	Muita	Media	Pouca	Cci	Muita	pouca
201	Muita	Media	Pouca	Ali	Pouca	pouca
202	Muita	Media	Pouca	Ali	Muita	pouca
203	Muita	Media	Pouca	Ccc	Pouca	pouca
204	Muita	Media	Pouca	Ccc	Muita	pouca
205	Muita	Media	Pouca	Chb	Pouca	pouca
206	Muita	Media	Pouca	Chb	Muita	pouca
207	Muita	Media	Pouca	Chi	Pouca	pouca
208	Muita	Media	Pouca	Chi	Muita	pouca
209	Muita	Media	Pouca	Hll	Pouca	pouca
210	Muita	Media	Pouca	Hll	Muita	pouca
211	Muita	Media	Muita	Cir	Pouca	pouca
212	Muita	Media	Muita	Cir	Muita	pouca
213	Muita	Media	Muita	Cci	Pouca	pouca
214	Muita	Media	Muita	Cci	Muita	pouca
215	Muita	Media	Muita	Ali	Pouca	pouca
216	Muita	Media	Muita	Ali	Muita	pouca
217	Muita	Media	Muita	Ccc	Pouca	pouca
218	Muita	Media	Muita	Ccc	Muita	pouca
219	Muita	Media	Muita	Chb	Pouca	pouca
220	Muita	Media	Muita	Chb	Muita	pouca
221	Muita	Media	Muita	Chi	Pouca	pouca
222	Muita	Media	Muita	Chi	Muita	pouca
223	Muita	Media	Muita	Hll	Pouca	pouca
224	Muita	Media	Muita	Hll	Muita	pouca
225	Muita	Muita	Pouca	Cir	Pouca	pouca
226	Muita	Muita	Pouca	Cir	Muita	pouca
227	Muita	Muita	Pouca	Cci	Pouca	pouca
228	Muita	Muita	Pouca	Cci	Muita	pouca
229	Muita	Muita	Pouca	Ali	Pouca	pouca
230	Muita	Muita	Pouca	Ali	Muita	pouca
231	Muita	Muita	Pouca	Ccc	Pouca	pouca
232	Muita	Muita	Pouca	Ccc	Muita	pouca
233	Muita	Muita	Pouca	Chb	Pouca	pouca
234	Muita	Muita	Pouca	Chb	Muita	pouca
235	Muita	Muita	Pouca	Chi	Pouca	pouca
236	Muita	Muita	Pouca	Chi	Muita	pouca
237	Muita	Muita	Pouca	Hll	Pouca	pouca
238	Muita	Muita	Pouca	Hll	Muita	pouca
239	Muita	Muita	Muita	Cir	Pouca	pouca

REGRA	CINZA	TAMANHO	ORIENTAÇÃO	FORMA	DISTÂNCIA	SALIÊNCIA TOPDOWN
240	Muita	Muita	Muita	Cir	Muita	pouca
241	Muita	Muita	Muita	Cci	Pouca	pouca
242	Muita	Muita	Muita	Cci	Muita	pouca
243	Muita	Muita	Muita	Ali	Pouca	pouca
244	Muita	Muita	Muita	Ali	Muita	pouca
245	Muita	Muita	Muita	Ccc	Pouca	pouca
246	Muita	Muita	Muita	Ccc	Muita	pouca
247	Muita	Muita	Muita	Chb	Pouca	pouca
248	Muita	Muita	Muita	Chb	Muita	pouca
249	Muita	Muita	Muita	Chi	Pouca	pouca
250	Muita	Muita	Muita	Chi	Muita	pouca
251	Muita	Muita	Muita	Hll	Pouca	pouca
252	Muita	Muita	Muita	Hll	Muita	pouca

Tabela A4.2 – Tabela de Regras para a tarefa de determinar os objetos “Grandes”

REGRA	CINZA	TAMANHO	ORIENTAÇÃO	FORMA	DISTÂNCIA	SALIÊNCIA TOPDOWN
1	Pouca	Pouca	Pouca	Cir	Pouca	pouca
2	Pouca	Pouca	Pouca	Cir	Muita	pouca
3	Pouca	Pouca	Pouca	Cci	Pouca	pouca
4	Pouca	Pouca	Pouca	Cci	Muita	pouca
5	Pouca	Pouca	Pouca	Ali	Pouca	pouca
6	Pouca	Pouca	Pouca	Ali	Muita	pouca
7	Pouca	Pouca	Pouca	Ccc	Pouca	pouca
8	Pouca	Pouca	Pouca	Ccc	Muita	pouca
9	Pouca	Pouca	Pouca	Chb	Pouca	pouca
10	Pouca	Pouca	Pouca	Chb	Muita	pouca
11	Pouca	Pouca	Pouca	Chi	Pouca	pouca
12	Pouca	Pouca	Pouca	Chi	Muita	pouca
13	Pouca	Pouca	Pouca	Hll	Pouca	pouca
14	Pouca	Pouca	Pouca	Hll	Muita	pouca
15	Pouca	Pouca	Muita	Cir	Pouca	pouca
16	Pouca	Pouca	Muita	Cir	Muita	pouca
17	Pouca	Pouca	Muita	Cci	Pouca	pouca
18	Pouca	Pouca	Muita	Cci	Muita	pouca
19	Pouca	Pouca	Muita	Ali	Pouca	pouca
20	Pouca	Pouca	Muita	Ali	Muita	pouca
21	Pouca	Pouca	Muita	Ccc	Pouca	pouca
22	Pouca	Pouca	Muita	Ccc	Muita	pouca
23	Pouca	Pouca	Muita	Chb	Pouca	pouca
24	Pouca	Pouca	Muita	Chb	Muita	pouca
25	Pouca	Pouca	Muita	Chi	Pouca	pouca
26	Pouca	Pouca	Muita	Chi	Muita	pouca
27	Pouca	Pouca	Muita	Hll	Pouca	pouca
28	Pouca	Pouca	Muita	Hll	Muita	pouca
29	Pouca	Média	Pouca	Cir	Pouca	media
30	Pouca	Media	Pouca	Cir	Muita	media

REGRA	CINZA	TAMANHO	ORIENTAÇÃO	FORMA	DISTÂNCIA	SALIÊNCIA TOPDOWN
31	Pouca	Media	Pouca	Cci	Pouca	media
32	Pouca	Media	Pouca	Cci	Muita	media
33	Pouca	Media	Pouca	Ali	Pouca	media
34	Pouca	Media	Pouca	Ali	Muita	media
35	Pouca	Media	Pouca	Ccc	Pouca	media
36	Pouca	Media	Pouca	Ccc	Muita	media
37	Pouca	Media	Pouca	Chb	Pouca	media
38	Pouca	Media	Pouca	Chb	Muita	media
39	Pouca	Media	Pouca	Chi	Pouca	media
40	Pouca	Media	Pouca	Chi	Muita	media
41	Pouca	Media	Pouca	Hll	Pouca	media
42	Pouca	Media	Pouca	Hll	Muita	media
43	Pouca	Media	Muita	Cir	Pouca	media
44	Pouca	Media	Muita	Cir	Muita	media
45	Pouca	Media	Muita	Cci	Pouca	media
46	Pouca	Media	Muita	Cci	Muita	media
47	Pouca	Media	Muita	Ali	Pouca	media
48	Pouca	Media	Muita	Ali	Muita	media
49	Pouca	Media	Muita	Ccc	Pouca	media
50	Pouca	Media	Muita	Ccc	Muita	media
51	Pouca	Media	Muita	Chb	Pouca	media
52	Pouca	Media	Muita	Chb	Muita	media
53	Pouca	Media	Muita	Chi	Pouca	media
54	Pouca	Media	Muita	Chi	Muita	media
55	Pouca	Media	Muita	Hll	Pouca	media
56	Pouca	Media	Muita	Hll	Muita	media
57	Pouca	Muita	Pouca	Cir	Pouca	muita
58	Pouca	Muita	Pouca	Cir	Muita	muita
59	Pouca	Muita	Pouca	Cci	Pouca	muita
60	Pouca	Muita	Pouca	Cci	Muita	muita
61	Pouca	Muita	Pouca	Ali	Pouca	muita
62	Pouca	Muita	Pouca	Ali	Muita	muita
63	Pouca	Muita	Pouca	Ccc	Pouca	muita
64	Pouca	Muita	Pouca	Ccc	Muita	muita
65	Pouca	Muita	Pouca	Chb	Pouca	muita
66	Pouca	Muita	Pouca	Chb	Muita	muita
67	Pouca	Muita	Pouca	Chi	Pouca	muita
68	Pouca	Muita	Pouca	Chi	Muita	muita
69	Pouca	Muita	Pouca	Hll	Pouca	muita
70	Pouca	Muita	Pouca	Hll	Muita	muita
71	Pouca	Muita	Muita	Cir	Pouca	muita
72	Pouca	Muita	Muita	Cir	Muita	muita
73	Pouca	Muita	Muita	Cci	Pouca	muita
74	Pouca	Muita	Muita	Cci	Muita	muita
75	Pouca	Muita	Muita	Ali	Pouca	muita
76	Pouca	Muita	Muita	Ali	Muita	muita
77	Pouca	Muita	Muita	Ccc	Pouca	muita
78	Pouca	Muita	Muita	Ccc	Muita	muita
79	Pouca	Muita	Muita	Chb	Pouca	muita
80	Pouca	Muita	Muita	Chb	Muita	muita

REGRA	CINZA	TAMANHO	ORIENTAÇÃO	FORMA	DISTÂNCIA	SALIÊNCIA TOPDOWN
81	Pouca	Muita	Muita	Chi	Pouca	muita
82	Pouca	Muita	Muita	Chi	Muita	muita
83	Pouca	Muita	Muita	Hll	Pouca	muita
84	Pouca	Muita	Muita	Hll	Muita	muita
85	Média	Pouca	Pouca	Cir	Pouca	pouca
86	Media	Pouca	Pouca	Cir	Muita	pouca
87	Media	Pouca	Pouca	Cci	Pouca	pouca
88	Media	Pouca	Pouca	Cci	Muita	pouca
89	Media	Pouca	Pouca	Ali	Pouca	pouca
90	Media	Pouca	Pouca	Ali	Muita	pouca
91	Media	Pouca	Pouca	Ccc	Pouca	pouca
92	Media	Pouca	Pouca	Ccc	Muita	pouca
93	Media	Pouca	Pouca	Chb	Pouca	pouca
94	Media	Pouca	Pouca	Chb	Muita	pouca
95	Media	Pouca	Pouca	Chi	Pouca	pouca
96	Media	Pouca	Pouca	Chi	Muita	pouca
97	Media	Pouca	Pouca	Hll	Pouca	pouca
98	Media	Pouca	Pouca	Hll	Muita	pouca
99	Media	Pouca	Muita	Cir	Pouca	pouca
100	Media	Pouca	Muita	Cir	Muita	pouca
101	Media	Pouca	Muita	Cci	Pouca	pouca
102	Media	Pouca	Muita	Cci	Muita	pouca
103	Media	Pouca	Muita	Ali	Pouca	pouca
104	Media	Pouca	Muita	Ali	Muita	pouca
105	Media	Pouca	Muita	Ccc	Pouca	pouca
106	Media	Pouca	Muita	Ccc	Muita	pouca
107	Media	Pouca	Muita	Chb	Pouca	pouca
108	Media	Pouca	Muita	Chb	Muita	pouca
109	Media	Pouca	Muita	Chi	Pouca	pouca
110	Media	Pouca	Muita	Chi	Muita	pouca
111	Media	Pouca	Muita	Hll	Pouca	pouca
112	Media	Pouca	Muita	Hll	Muita	pouca
113	Média	Média	Pouca	Cir	Pouca	media
114	Media	Media	Pouca	Cir	Muita	media
115	Media	Media	Pouca	Cci	Pouca	media
116	Media	Media	Pouca	Cci	Muita	media
117	Media	Media	Pouca	Ali	Pouca	media
118	Media	Media	Pouca	Ali	Muita	media
119	Media	Media	Pouca	Ccc	Pouca	media
120	Media	Media	Pouca	Ccc	Muita	media
121	Media	Media	Pouca	Chb	Pouca	media
122	Media	Media	Pouca	Chb	Muita	media
123	Media	Media	Pouca	Chi	Pouca	media
124	Media	Media	Pouca	Chi	Muita	media
124	Media	Media	Pouca	Hll	Pouca	media
126	Media	Media	Pouca	Hll	Muita	media
127	Media	Media	Muita	Cir	Pouca	media
128	Media	Media	Muita	Cir	Muita	media
129	Media	Media	Muita	Cci	Pouca	media
130	Media	Media	Muita	Cci	Muita	media

REGRA	CINZA	TAMANHO	ORIENTAÇÃO	FORMA	DISTÂNCIA	SALIÊNCIA TOPDOWN
131	Media	Media	Muita	Ali	Pouca	media
132	Media	Media	Muita	Ali	Muita	media
133	Media	Media	Muita	Ccc	Pouca	media
134	Media	Media	Muita	Ccc	Muita	media
135	Media	Media	Muita	Chb	Pouca	media
136	Media	Media	Muita	Chb	Muita	media
137	Media	Media	Muita	Chi	Pouca	media
138	Media	Media	Muita	Chi	Muita	media
139	Media	Media	Muita	Hll	Pouca	media
140	Media	Media	Muita	Hll	Muita	media
141	Media	Muita	Pouca	Cir	Pouca	muita
142	Media	Muita	Pouca	Cir	Muita	muita
143	Media	Muita	Pouca	Cci	Pouca	muita
144	Media	Muita	Pouca	Cci	Muita	muita
145	Media	Muita	Pouca	Ali	Pouca	muita
146	Media	Muita	Pouca	Ali	Muita	muita
147	Media	Muita	Pouca	Ccc	Pouca	muita
148	Media	Muita	Pouca	Ccc	Muita	muita
149	Media	Muita	Pouca	Chb	Pouca	muita
150	Media	Muita	Pouca	Chb	Muita	muita
151	Media	Muita	Pouca	Chi	Pouca	muita
152	Media	Muita	Pouca	Chi	Muita	muita
153	Media	Muita	Pouca	Hll	Pouca	muita
154	Media	Muita	Pouca	Hll	Muita	muita
155	Media	Muita	Muita	Cir	Pouca	muita
156	Media	Muita	Muita	Cir	Muita	muita
157	Media	Muita	Muita	Cci	Pouca	muita
158	Media	Muita	Muita	Cci	Muita	muita
159	Media	Muita	Muita	Ali	Pouca	muita
160	Media	Muita	Muita	Ali	Muita	muita
161	Media	Muita	Muita	Ccc	Pouca	muita
162	Media	Muita	Muita	Ccc	Muita	muita
163	Media	Muita	Muita	Chb	Pouca	muita
164	Media	Muita	Muita	Chb	Muita	muita
165	Media	Muita	Muita	Chi	Pouca	muita
166	Media	Muita	Muita	Chi	Muita	muita
167	Media	Muita	Muita	Hll	Pouca	muita
168	Media	Muita	Muita	Hll	Muita	muita
169	Muita	Pouca	Pouca	Cir	Pouca	pouca
170	Muita	Pouca	Pouca	Cir	Muita	pouca
171	Muita	Pouca	Pouca	Cci	Pouca	pouca
172	Muita	Pouca	Pouca	Cci	Muita	pouca
173	Muita	Pouca	Pouca	Ali	Pouca	pouca
174	Muita	Pouca	Pouca	Ali	Muita	pouca
175	Muita	Pouca	Pouca	Ccc	Pouca	pouca
176	Muita	Pouca	Pouca	Ccc	Muita	pouca
177	Muita	Pouca	Pouca	Chb	Pouca	pouca
178	Muita	Pouca	Pouca	Chb	Muita	pouca
179	Muita	Pouca	Pouca	Chi	Pouca	pouca
180	Muita	Pouca	Pouca	Chi	Muita	pouca

REGRA	CINZA	TAMANHO	ORIENTAÇÃO	FORMA	DISTÂNCIA	SALIÊNCIA TOPDOWN
181	Muita	Pouca	Pouca	Hll	Pouca	pouca
182	Muita	Pouca	Pouca	Hll	Muita	pouca
183	Muita	Pouca	Muita	Cir	Pouca	pouca
184	Muita	Pouca	Muita	Cir	Muita	pouca
185	Muita	Pouca	Muita	Cci	Pouca	pouca
186	Muita	Pouca	Muita	Cci	Muita	pouca
187	Muita	Pouca	Muita	Ali	Pouca	pouca
188	Muita	Pouca	Muita	Ali	Muita	pouca
189	Muita	Pouca	Muita	Ccc	Pouca	pouca
190	Muita	Pouca	Muita	Ccc	Muita	pouca
191	Muita	Pouca	Muita	Chb	Pouca	pouca
192	Muita	Pouca	Muita	Chb	Muita	pouca
193	Muita	Pouca	Muita	Chi	Pouca	pouca
194	Muita	Pouca	Muita	Chi	Muita	pouca
195	Muita	Pouca	Muita	Hll	Pouca	pouca
196	Muita	Pouca	Muita	Hll	Muita	pouca
197	Muita	Média	Pouca	Cir	Pouca	media
198	Muita	Media	Pouca	Cir	Muita	media
199	Muita	Media	Pouca	Cci	Pouca	media
200	Muita	Media	Pouca	Cci	Muita	media
201	Muita	Media	Pouca	Ali	Pouca	media
202	Muita	Media	Pouca	Ali	Muita	media
203	Muita	Media	Pouca	Ccc	Pouca	media
204	Muita	Media	Pouca	Ccc	Muita	media
205	Muita	Media	Pouca	Chb	Pouca	media
206	Muita	Media	Pouca	Chb	Muita	media
207	Muita	Media	Pouca	Chi	Pouca	media
208	Muita	Media	Pouca	Chi	Muita	media
209	Muita	Media	Pouca	Hll	Pouca	media
210	Muita	Media	Pouca	Hll	Muita	media
211	Muita	Media	Muita	Cir	Pouca	media
212	Muita	Media	Muita	Cir	Muita	media
213	Muita	Media	Muita	Cci	Pouca	media
214	Muita	Media	Muita	Cci	Muita	media
215	Muita	Media	Muita	Ali	Pouca	media
216	Muita	Media	Muita	Ali	Muita	media
217	Muita	Media	Muita	Ccc	Pouca	media
218	Muita	Media	Muita	Ccc	Muita	media
219	Muita	Media	Muita	Chb	Pouca	media
220	Muita	Media	Muita	Chb	Muita	media
221	Muita	Media	Muita	Chi	Pouca	media
222	Muita	Media	Muita	Chi	Muita	media
223	Muita	Media	Muita	Hll	Pouca	media
224	Muita	Media	Muita	Hll	Muita	media
225	Muita	Muita	Pouca	Cir	Pouca	muita
226	Muita	Muita	Pouca	Cir	Muita	muita
227	Muita	Muita	Pouca	Cci	Pouca	muita
228	Muita	Muita	Pouca	Cci	Muita	muita
229	Muita	Muita	Pouca	Ali	Pouca	muita
230	Muita	Muita	Pouca	Ali	Muita	muita

REGRA	CINZA	TAMANHO	ORIENTAÇÃO	FORMA	DISTÂNCIA	SALIÊNCIA TOPDOWN
231	Muita	Muita	Pouca	Ccc	Pouca	muita
232	Muita	Muita	Pouca	Ccc	Muita	muita
233	Muita	Muita	Pouca	Chb	Pouca	muita
234	Muita	Muita	Pouca	Chb	Muita	muita
235	Muita	Muita	Pouca	Chi	Pouca	muita
236	Muita	Muita	Pouca	Chi	Muita	muita
237	Muita	Muita	Pouca	Hll	Pouca	muita
238	Muita	Muita	Pouca	Hll	Muita	muita
239	Muita	Muita	Muita	Cir	Pouca	muita
240	Muita	Muita	Muita	Cir	Muita	muita
241	Muita	Muita	Muita	Cci	Pouca	muita
242	Muita	Muita	Muita	Cci	Muita	muita
243	Muita	Muita	Muita	Ali	Pouca	muita
244	Muita	Muita	Muita	Ali	Muita	muita
245	Muita	Muita	Muita	Ccc	Pouca	muita
246	Muita	Muita	Muita	Ccc	Muita	muita
247	Muita	Muita	Muita	Chb	Pouca	muita
248	Muita	Muita	Muita	Chb	Muita	muita
249	Muita	Muita	Muita	Chi	Pouca	muita
250	Muita	Muita	Muita	Chi	Muita	muita
251	Muita	Muita	Muita	Hll	Pouca	muita
252	Muita	Muita	Muita	Hll	Muita	muita

Tabela A4.3 – Regras para a tarefa de determinar os Círculos

REGRA	CINZA	TAMANHO	ORIENTAÇÃO	FORMA	DISTÂNCIA	SALIÊNCIA TOPDOWN
1	Pouca	Pouca	Pouca	Cir	Pouca	muita
2	Pouca	Pouca	Pouca	Cir	Muita	muita
3	Pouca	Pouca	Pouca	Cci	Pouca	pouca
4	Pouca	Pouca	Pouca	Cci	Muita	pouca
5	Pouca	Pouca	Pouca	Ali	Pouca	pouca
6	Pouca	Pouca	Pouca	Ali	Muita	pouca
7	Pouca	Pouca	Pouca	Ccc	Pouca	pouca
8	Pouca	Pouca	Pouca	Ccc	Muita	pouca
9	Pouca	Pouca	Pouca	Chb	Pouca	pouca
10	Pouca	Pouca	Pouca	Chb	Muita	pouca
11	Pouca	Pouca	Pouca	Chi	Pouca	pouca
12	Pouca	Pouca	Pouca	Chi	Muita	pouca
13	Pouca	Pouca	Pouca	Hll	Pouca	pouca
14	Pouca	Pouca	Pouca	Hll	Muita	pouca
15	Pouca	Pouca	Muita	Cir	Pouca	muita
16	Pouca	Pouca	Muita	Cir	Muita	muita
17	Pouca	Pouca	Muita	Cci	Pouca	pouca
18	Pouca	Pouca	Muita	Cci	Muita	pouca
19	Pouca	Pouca	Muita	Ali	Pouca	pouca
20	Pouca	Pouca	Muita	Ali	Muita	pouca
21	Pouca	Pouca	Muita	Ccc	Pouca	pouca

REGRA	CINZA	TAMANHO	ORIENTAÇÃO	FORMA	DISTÂNCIA	SALIÊNCIA TOPDOWN
22	Pouca	Pouca	Muita	Ccc	Muita	pouca
23	Pouca	Pouca	Muita	Chb	Pouca	pouca
24	Pouca	Pouca	Muita	Chb	Muita	pouca
25	Pouca	Pouca	Muita	Chi	Pouca	pouca
26	Pouca	Pouca	Muita	Chi	Muita	pouca
27	Pouca	Pouca	Muita	Hll	Pouca	pouca
28	Pouca	Pouca	Muita	Hll	Muita	pouca
29	Pouca	Média	Pouca	Cir	Pouca	muita
30	Pouca	Media	Pouca	Cir	Muita	muita
31	Pouca	Media	Pouca	Cci	Pouca	pouca
32	Pouca	Media	Pouca	Cci	Muita	pouca
33	Pouca	Media	Pouca	Ali	Pouca	pouca
34	Pouca	Media	Pouca	Ali	Muita	pouca
35	Pouca	Media	Pouca	Ccc	Pouca	pouca
36	Pouca	Media	Pouca	Ccc	Muita	pouca
37	Pouca	Media	Pouca	Chb	Pouca	pouca
38	Pouca	Media	Pouca	Chb	Muita	pouca
39	Pouca	Media	Pouca	Chi	Pouca	pouca
40	Pouca	Media	Pouca	Chi	Muita	pouca
41	Pouca	Media	Pouca	Hll	Pouca	pouca
42	Pouca	Media	Pouca	Hll	Muita	pouca
43	Pouca	Media	Muita	Cir	Pouca	muita
44	Pouca	Media	Muita	Cir	Muita	muita
45	Pouca	Media	Muita	Cci	Pouca	pouca
46	Pouca	Media	Muita	Cci	Muita	pouca
47	Pouca	Media	Muita	Ali	Pouca	pouca
48	Pouca	Media	Muita	Ali	Muita	pouca
49	Pouca	Media	Muita	Ccc	Pouca	pouca
50	Pouca	Media	Muita	Ccc	Muita	pouca
51	Pouca	Media	Muita	Chb	Pouca	pouca
52	Pouca	Media	Muita	Chb	Muita	pouca
53	Pouca	Media	Muita	Chi	Pouca	pouca
54	Pouca	Media	Muita	Chi	Muita	pouca
55	Pouca	Media	Muita	Hll	Pouca	pouca
56	Pouca	Media	Muita	Hll	Muita	pouca
57	Pouca	Muita	Pouca	Cir	Pouca	muita
58	Pouca	Muita	Pouca	Cir	Muita	muita
59	Pouca	Muita	Pouca	Cci	Pouca	pouca
60	Pouca	Muita	Pouca	Cci	Muita	pouca
61	Pouca	Muita	Pouca	Ali	Pouca	pouca
62	Pouca	Muita	Pouca	Ali	Muita	pouca
63	Pouca	Muita	Pouca	Ccc	Pouca	pouca
64	Pouca	Muita	Pouca	Ccc	Muita	pouca
65	Pouca	Muita	Pouca	Chb	Pouca	pouca
66	Pouca	Muita	Pouca	Chb	Muita	pouca
67	Pouca	Muita	Pouca	Chi	Pouca	pouca
68	Pouca	Muita	Pouca	Chi	Muita	pouca
69	Pouca	Muita	Pouca	Hll	Pouca	pouca
70	Pouca	Muita	Pouca	Hll	Muita	pouca
71	Pouca	Muita	Muita	Cir	Pouca	muita

REGRA	CINZA	TAMANHO	ORIENTAÇÃO	FORMA	DISTÂNCIA	SALIÊNCIA TOPDOWN
72	Pouca	Muita	Muita	Cir	Muita	muita
73	Pouca	Muita	Muita	Cci	Pouca	pouca
74	Pouca	Muita	Muita	Cci	Muita	pouca
75	Pouca	Muita	Muita	Ali	Pouca	pouca
76	Pouca	Muita	Muita	Ali	Muita	pouca
77	Pouca	Muita	Muita	Ccc	Pouca	pouca
78	Pouca	Muita	Muita	Ccc	Muita	pouca
79	Pouca	Muita	Muita	Chb	Pouca	pouca
80	Pouca	Muita	Muita	Chb	Muita	pouca
81	Pouca	Muita	Muita	Chi	Pouca	pouca
82	Pouca	Muita	Muita	Chi	Muita	pouca
83	Pouca	Muita	Muita	Hll	Pouca	pouca
84	Pouca	Muita	Muita	Hll	Muita	pouca
85	Média	Pouca	Pouca	Cir	Pouca	muita
86	Media	Pouca	Pouca	Cir	Muita	muita
87	Media	Pouca	Pouca	Cci	Pouca	pouca
88	Media	Pouca	Pouca	Cci	Muita	pouca
89	Media	Pouca	Pouca	Ali	Pouca	pouca
90	Media	Pouca	Pouca	Ali	Muita	pouca
91	Media	Pouca	Pouca	Ccc	Pouca	pouca
92	Media	Pouca	Pouca	Ccc	Muita	pouca
93	Media	Pouca	Pouca	Chb	Pouca	pouca
94	Media	Pouca	Pouca	Chb	Muita	pouca
95	Media	Pouca	Pouca	Chi	Pouca	pouca
96	Media	Pouca	Pouca	Chi	Muita	pouca
97	Media	Pouca	Pouca	Hll	Pouca	pouca
98	Media	Pouca	Pouca	Hll	Muita	pouca
99	Media	Pouca	Muita	Cir	Pouca	muita
100	Media	Pouca	Muita	Cir	Muita	muita
101	Media	Pouca	Muita	Cci	Pouca	pouca
102	Media	Pouca	Muita	Cci	Muita	pouca
103	Media	Pouca	Muita	Ali	Pouca	pouca
104	Media	Pouca	Muita	Ali	Muita	pouca
105	Media	Pouca	Muita	Ccc	Pouca	pouca
106	Media	Pouca	Muita	Ccc	Muita	pouca
107	Media	Pouca	Muita	Chb	Pouca	pouca
108	Media	Pouca	Muita	Chb	Muita	pouca
109	Media	Pouca	Muita	Chi	Pouca	pouca
110	Media	Pouca	Muita	Chi	Muita	pouca
111	Media	Pouca	Muita	Hll	Pouca	pouca
112	Media	Pouca	Muita	Hll	Muita	pouca
113	Média	Média	Pouca	Cir	Pouca	muita
114	Media	Media	Pouca	Cir	Muita	muita
115	Media	Media	Pouca	Cci	Pouca	pouca
116	Media	Media	Pouca	Cci	Muita	pouca
117	Media	Media	Pouca	Ali	Pouca	pouca
118	Media	Media	Pouca	Ali	Muita	pouca
119	Media	Media	Pouca	Ccc	Pouca	pouca
120	Media	Media	Pouca	Ccc	Muita	pouca
121	Media	Media	Pouca	Chb	Pouca	pouca

REGRA	CINZA	TAMANHO	ORIENTAÇÃO	FORMA	DISTÂNCIA	SALIÊNCIA TOPDOWN
122	Media	Media	Pouca	Chb	Muita	pouca
123	Media	Media	Pouca	Chi	Pouca	pouca
124	Media	Media	Pouca	Chi	Muita	pouca
124	Media	Media	Pouca	Hll	Pouca	pouca
126	Media	Media	Pouca	Hll	Muita	pouca
127	Media	Media	Muita	Cir	Pouca	muita
128	Media	Media	Muita	Cir	Muita	muita
129	Media	Media	Muita	Cci	Pouca	pouca
130	Media	Media	Muita	Cci	Muita	pouca
131	Media	Media	Muita	Ali	Pouca	pouca
132	Media	Media	Muita	Ali	Muita	pouca
133	Media	Media	Muita	Ccc	Pouca	pouca
134	Media	Media	Muita	Ccc	Muita	pouca
135	Media	Media	Muita	Chb	Pouca	pouca
136	Media	Media	Muita	Chb	Muita	pouca
137	Media	Media	Muita	Chi	Pouca	pouca
138	Media	Media	Muita	Chi	Muita	pouca
139	Media	Media	Muita	Hll	Pouca	pouca
140	Media	Media	Muita	Hll	Muita	pouca
141	Media	Muita	Pouca	Cir	Pouca	muita
142	Media	Muita	Pouca	Cir	Muita	muita
143	Media	Muita	Pouca	Cci	Pouca	pouca
144	Media	Muita	Pouca	Cci	Muita	pouca
145	Media	Muita	Pouca	Ali	Pouca	pouca
146	Media	Muita	Pouca	Ali	Muita	pouca
147	Media	Muita	Pouca	Ccc	Pouca	pouca
148	Media	Muita	Pouca	Ccc	Muita	pouca
149	Media	Muita	Pouca	Chb	Pouca	pouca
150	Media	Muita	Pouca	Chb	Muita	pouca
151	Media	Muita	Pouca	Chi	Pouca	pouca
152	Media	Muita	Pouca	Chi	Muita	pouca
153	Media	Muita	Pouca	Hll	Pouca	pouca
154	Media	Muita	Pouca	Hll	Muita	pouca
155	Media	Muita	Muita	Cir	Pouca	muita
156	Media	Muita	Muita	Cir	Muita	muita
157	Media	Muita	Muita	Cci	Pouca	pouca
158	Media	Muita	Muita	Cci	Muita	pouca
159	Media	Muita	Muita	Ali	Pouca	pouca
160	Media	Muita	Muita	Ali	Muita	pouca
161	Media	Muita	Muita	Ccc	Pouca	pouca
162	Media	Muita	Muita	Ccc	Muita	pouca
163	Media	Muita	Muita	Chb	Pouca	pouca
164	Media	Muita	Muita	Chb	Muita	pouca
165	Media	Muita	Muita	Chi	Pouca	pouca
166	Media	Muita	Muita	Chi	Muita	pouca
167	Media	Muita	Muita	Hll	Pouca	pouca
168	Media	Muita	Muita	Hll	Muita	pouca
169	Muita	Pouca	Pouca	Cir	Pouca	muita
170	Muita	Pouca	Pouca	Cir	Muita	muita
171	Muita	Pouca	Pouca	Cci	Pouca	pouca

REGRA	CINZA	TAMANHO	ORIENTAÇÃO	FORMA	DISTÂNCIA	SALIÊNCIA TOPDOWN
172	Muita	Pouca	Pouca	Cci	Muita	pouca
173	Muita	Pouca	Pouca	Ali	Pouca	pouca
174	Muita	Pouca	Pouca	Ali	Muita	pouca
175	Muita	Pouca	Pouca	Ccc	Pouca	pouca
176	Muita	Pouca	Pouca	Ccc	Muita	pouca
177	Muita	Pouca	Pouca	Chb	Pouca	pouca
178	Muita	Pouca	Pouca	Chb	Muita	pouca
179	Muita	Pouca	Pouca	Chi	Pouca	pouca
180	Muita	Pouca	Pouca	Chi	Muita	pouca
181	Muita	Pouca	Pouca	Hll	Pouca	pouca
182	Muita	Pouca	Pouca	Hll	Muita	pouca
183	Muita	Pouca	Muita	Cir	Pouca	muita
184	Muita	Pouca	Muita	Cir	Muita	muita
185	Muita	Pouca	Muita	Cci	Pouca	pouca
186	Muita	Pouca	Muita	Cci	Muita	pouca
187	Muita	Pouca	Muita	Ali	Pouca	pouca
188	Muita	Pouca	Muita	Ali	Muita	pouca
189	Muita	Pouca	Muita	Ccc	Pouca	pouca
190	Muita	Pouca	Muita	Ccc	Muita	pouca
191	Muita	Pouca	Muita	Chb	Pouca	pouca
192	Muita	Pouca	Muita	Chb	Muita	pouca
193	Muita	Pouca	Muita	Chi	Pouca	pouca
194	Muita	Pouca	Muita	Chi	Muita	pouca
195	Muita	Pouca	Muita	Hll	Pouca	pouca
196	Muita	Pouca	Muita	Hll	Muita	pouca
197	Muita	Média	Pouca	Cir	Pouca	muita
198	Muita	Media	Pouca	Cir	Muita	muita
199	Muita	Media	Pouca	Cci	Pouca	pouca
200	Muita	Media	Pouca	Cci	Muita	pouca
201	Muita	Media	Pouca	Ali	Pouca	pouca
202	Muita	Media	Pouca	Ali	Muita	pouca
203	Muita	Media	Pouca	Ccc	Pouca	pouca
204	Muita	Media	Pouca	Ccc	Muita	pouca
205	Muita	Media	Pouca	Chb	Pouca	pouca
206	Muita	Media	Pouca	Chb	Muita	pouca
207	Muita	Media	Pouca	Chi	Pouca	pouca
208	Muita	Media	Pouca	Chi	Muita	pouca
209	Muita	Media	Pouca	Hll	Pouca	pouca
210	Muita	Media	Pouca	Hll	Muita	pouca
211	Muita	Media	Muita	Cir	Pouca	muita
212	Muita	Media	Muita	Cir	Muita	muita
213	Muita	Media	Muita	Cci	Pouca	pouca
214	Muita	Media	Muita	Cci	Muita	pouca
215	Muita	Media	Muita	Ali	Pouca	pouca
216	Muita	Media	Muita	Ali	Muita	pouca
217	Muita	Media	Muita	Ccc	Pouca	pouca
218	Muita	Media	Muita	Ccc	Muita	pouca
219	Muita	Media	Muita	Chb	Pouca	pouca
220	Muita	Media	Muita	Chb	Muita	pouca
221	Muita	Media	Muita	Chi	Pouca	pouca

REGRA	CINZA	TAMANHO	ORIENTAÇÃO	FORMA	DISTÂNCIA	SALIÊNCIA TOPDOWN
222	Muita	Media	Muita	Chi	Muita	pouca
223	Muita	Media	Muita	Hll	Pouca	pouca
224	Muita	Media	Muita	Hll	Muita	pouca
225	Muita	Muita	Pouca	Cir	Pouca	muita
226	Muita	Muita	Pouca	Cir	Muita	muita
227	Muita	Muita	Pouca	Cci	Pouca	pouca
228	Muita	Muita	Pouca	Cci	Muita	pouca
229	Muita	Muita	Pouca	Ali	Pouca	pouca
230	Muita	Muita	Pouca	Ali	Muita	pouca
231	Muita	Muita	Pouca	Ccc	Pouca	pouca
232	Muita	Muita	Pouca	Ccc	Muita	pouca
233	Muita	Muita	Pouca	Chb	Pouca	pouca
234	Muita	Muita	Pouca	Chb	Muita	pouca
235	Muita	Muita	Pouca	Chi	Pouca	pouca
236	Muita	Muita	Pouca	Chi	Muita	pouca
237	Muita	Muita	Pouca	Hll	Pouca	pouca
238	Muita	Muita	Pouca	Hll	Muita	pouca
239	Muita	Muita	Muita	Cir	Pouca	muita
240	Muita	Muita	Muita	Cir	Muita	muita
241	Muita	Muita	Muita	Cci	Pouca	pouca
242	Muita	Muita	Muita	Cci	Muita	pouca
243	Muita	Muita	Muita	Ali	Pouca	pouca
244	Muita	Muita	Muita	Ali	Muita	pouca
245	Muita	Muita	Muita	Ccc	Pouca	pouca
246	Muita	Muita	Muita	Ccc	Muita	pouca
247	Muita	Muita	Muita	Chb	Pouca	pouca
248	Muita	Muita	Muita	Chb	Muita	pouca
249	Muita	Muita	Muita	Chi	Pouca	pouca
250	Muita	Muita	Muita	Chi	Muita	pouca
251	Muita	Muita	Muita	Hll	Pouca	pouca
252	Muita	Muita	Muita	Hll	Muita	pouca

Tabela A4.4 – Tabela de Regras para a tarefa de determinar os “Círculos Grandes”

REGRA	CINZA	TAMANHO	ORIENTAÇÃO	FORMA	DISTÂNCIA	SALIÊNCIA TOPDOWN
1	Pouca	Pouca	Pouca	Cir	Pouca	Media
2	Pouca	Pouca	Pouca	Cir	Muita	Media
3	Pouca	Pouca	Pouca	Cci	Pouca	Pouca
4	Pouca	Pouca	Pouca	Cci	Muita	Pouca
5	Pouca	Pouca	Pouca	Ali	Pouca	Pouca
6	Pouca	Pouca	Pouca	Ali	Muita	Pouca
7	Pouca	Pouca	Pouca	Ccc	Pouca	Pouca
8	Pouca	Pouca	Pouca	Ccc	Muita	Pouca
9	Pouca	Pouca	Pouca	Chb	Pouca	Pouca
10	Pouca	Pouca	Pouca	Chb	Muita	Pouca
11	Pouca	Pouca	Pouca	Chi	Pouca	Pouca
12	Pouca	Pouca	Pouca	Chi	Muita	Pouca

REGRA	CINZA	TAMANHO	ORIENTAÇÃO	FORMA	DISTANCIA	SALIÊNCIA TOPDOWN
13	Pouca	Pouca	Pouca	Hll	Pouca	Pouca
14	Pouca	Pouca	Pouca	Hll	Muita	Pouca
15	Pouca	Pouca	Muita	Cir	Pouca	Media
16	Pouca	Pouca	Muita	Cir	Muita	Media
17	Pouca	Pouca	Muita	Cci	Pouca	Pouca
18	Pouca	Pouca	Muita	Cci	Muita	Pouca
19	Pouca	Pouca	Muita	Ali	Pouca	Pouca
20	Pouca	Pouca	Muita	Ali	Muita	Pouca
21	Pouca	Pouca	Muita	Ccc	Pouca	Pouca
22	Pouca	Pouca	Muita	Ccc	Muita	Pouca
23	Pouca	Pouca	Muita	Chb	Pouca	Pouca
24	Pouca	Pouca	Muita	Chb	Muita	Pouca
25	Pouca	Pouca	Muita	Chi	Pouca	Pouca
26	Pouca	Pouca	Muita	Chi	Muita	Pouca
27	Pouca	Pouca	Muita	Hll	Pouca	Pouca
28	Pouca	Pouca	Muita	Hll	Muita	Pouca
29	Pouca	Média	Pouca	Cir	Pouca	Media-Muita
30	Pouca	Media	Pouca	Cir	Muita	Média-Muita
31	Pouca	Media	Pouca	Cci	Pouca	Pouca
32	Pouca	Media	Pouca	Cci	Muita	Pouca
33	Pouca	Media	Pouca	Ali	Pouca	Pouca
34	Pouca	Media	Pouca	Ali	Muita	Pouca
35	Pouca	Media	Pouca	Ccc	Pouca	Pouca
36	Pouca	Media	Pouca	Ccc	Muita	Pouca
37	Pouca	Media	Pouca	Chb	Pouca	Pouca
38	Pouca	Media	Pouca	Chb	Muita	Pouca
39	Pouca	Media	Pouca	Chi	Pouca	Pouca
40	Pouca	Media	Pouca	Chi	Muita	Pouca
41	Pouca	Media	Pouca	Hll	Pouca	Pouca
42	Pouca	Media	Pouca	Hll	Muita	Pouca
43	Pouca	Media	Muita	Cir	Pouca	Média-Muita
44	Pouca	Media	Muita	Cir	Muita	Média-Muita
45	Pouca	Media	Muita	Cci	Pouca	Pouca
46	Pouca	Media	Muita	Cci	Muita	Pouca
47	Pouca	Media	Muita	Ali	Pouca	Pouca
48	Pouca	Media	Muita	Ali	Muita	Pouca
49	Pouca	Media	Muita	Ccc	Pouca	Pouca
50	Pouca	Media	Muita	Ccc	Muita	Pouca
51	Pouca	Media	Muita	Chb	Pouca	Pouca
52	Pouca	Media	Muita	Chb	Muita	Pouca
53	Pouca	Media	Muita	Chi	Pouca	Pouca
54	Pouca	Media	Muita	Chi	Muita	Pouca
55	Pouca	Media	Muita	Hll	Pouca	Pouca
56	Pouca	Media	Muita	Hll	Muita	Pouca
57	Pouca	Muita	Pouca	Cir	Pouca	Muita
58	Pouca	Muita	Pouca	Cir	Muita	Muita
59	Pouca	Muita	Pouca	Cci	Pouca	Media-pouca
60	Pouca	Muita	Pouca	Cci	Muita	Media-pouca
61	Pouca	Muita	Pouca	Ali	Pouca	Media-pouca
62	Pouca	Muita	Pouca	Ali	Muita	Media-pouca

REGRA	CINZA	TAMANHO	ORIENTAÇÃO	FORMA	DISTÂNCIA	SALIÊNCIA TOPDOWN
63	Pouca	Muita	Pouca	Ccc	Pouca	Media-pouca
64	Pouca	Muita	Pouca	Ccc	Muita	Media-pouca
65	Pouca	Muita	Pouca	Chb	Pouca	Media-pouca
66	Pouca	Muita	Pouca	Chb	Muita	Media-pouca
67	Pouca	Muita	Pouca	Chi	Pouca	Media-pouca
68	Pouca	Muita	Pouca	Chi	Muita	Media-pouca
69	Pouca	Muita	Pouca	Hll	Pouca	Media-pouca
70	Pouca	Muita	Pouca	Hll	Muita	Media-pouca
71	Pouca	Muita	Muita	Cir	Pouca	Muita
72	Pouca	Muita	Muita	Cir	Muita	Muita
73	Pouca	Muita	Muita	Cci	Pouca	Media-pouca
74	Pouca	Muita	Muita	Cci	Muita	Media-pouca
75	Pouca	Muita	Muita	Ali	Pouca	Media-pouca
76	Pouca	Muita	Muita	Ali	Muita	Media-pouca
77	Pouca	Muita	Muita	Ccc	Pouca	Media-pouca
78	Pouca	Muita	Muita	Ccc	Muita	Media-pouca
79	Pouca	Muita	Muita	Chb	Pouca	Media-pouca
80	Pouca	Muita	Muita	Chb	Muita	Media-pouca
81	Pouca	Muita	Muita	Chi	Pouca	Media-pouca
82	Pouca	Muita	Muita	Chi	Muita	Media-pouca
83	Pouca	Muita	Muita	Hll	Pouca	Media-pouca
84	Pouca	Muita	Muita	Hll	Muita	Media-pouca
85	Média	Pouca	Pouca	Cir	Pouca	Media
86	Media	Pouca	Pouca	Cir	Muita	Media
87	Media	Pouca	Pouca	Cci	Pouca	Pouca
88	Media	Pouca	Pouca	Cci	Muita	Pouca
89	Media	Pouca	Pouca	Ali	Pouca	Pouca
90	Media	Pouca	Pouca	Ali	Muita	Pouca
91	Media	Pouca	Pouca	Ccc	Pouca	Pouca
92	Media	Pouca	Pouca	Ccc	Muita	Pouca
93	Media	Pouca	Pouca	Chb	Pouca	Pouca
94	Media	Pouca	Pouca	Chb	Muita	Pouca
95	Media	Pouca	Pouca	Chi	Pouca	Pouca
96	Media	Pouca	Pouca	Chi	Muita	Pouca
97	Media	Pouca	Pouca	Hll	Pouca	Pouca
98	Media	Pouca	Pouca	Hll	Muita	Pouca
99	Media	Pouca	Muita	Cir	Pouca	Media
100	Media	Pouca	Muita	Cir	Muita	Media
101	Media	Pouca	Muita	Cci	Pouca	Pouca
102	Media	Pouca	Muita	Cci	Muita	Pouca
103	Media	Pouca	Muita	Ali	Pouca	Pouca
104	Media	Pouca	Muita	Ali	Muita	Pouca
105	Media	Pouca	Muita	Ccc	Pouca	Pouca
106	Media	Pouca	Muita	Ccc	Muita	Pouca
107	Media	Pouca	Muita	Chb	Pouca	Pouca
108	Media	Pouca	Muita	Chb	Muita	Pouca
109	Media	Pouca	Muita	Chi	Pouca	Pouca
110	Media	Pouca	Muita	Chi	Muita	Pouca
111	Media	Pouca	Muita	Hll	Pouca	Pouca
112	Media	Pouca	Muita	Hll	Muita	Pouca

REGRA	CINZA	TAMANHO	ORIENTAÇÃO	FORMA	DISTÂNCIA	SALIÊNCIA TOPDOWN
113	Média	Média	Pouca	Cir	Pouca	Media-Muita
114	Media	Media	Pouca	Cir	Muita	Média-Muita
115	Media	Media	Pouca	Cci	Pouca	Pouca
116	Media	Media	Pouca	Cci	Muita	Pouca
117	Media	Media	Pouca	Ali	Pouca	Pouca
118	Media	Media	Pouca	Ali	Muita	Pouca
119	Media	Media	Pouca	Ccc	Pouca	Pouca
120	Media	Media	Pouca	Ccc	Muita	Pouca
121	Media	Media	Pouca	Chb	Pouca	Pouca
122	Media	Media	Pouca	Chb	Muita	Pouca
123	Media	Media	Pouca	Chi	Pouca	Pouca
124	Media	Media	Pouca	Chi	Muita	Pouca
124	Media	Media	Pouca	Hll	Pouca	Pouca
126	Media	Media	Pouca	Hll	Muita	Pouca
127	Media	Media	Muita	Cir	Pouca	Média-Muita
128	Media	Media	Muita	Cir	Muita	Média-Muita
129	Media	Media	Muita	Cci	Pouca	Pouca
130	Media	Media	Muita	Cci	Muita	Pouca
131	Media	Media	Muita	Ali	Pouca	Pouca
132	Media	Media	Muita	Ali	Muita	Pouca
133	Media	Media	Muita	Ccc	Pouca	Pouca
134	Media	Media	Muita	Ccc	Muita	Pouca
135	Media	Media	Muita	Chb	Pouca	Pouca
136	Media	Media	Muita	Chb	Muita	Pouca
137	Media	Media	Muita	Chi	Pouca	Pouca
138	Media	Media	Muita	Chi	Muita	Pouca
139	Media	Media	Muita	Hll	Pouca	Pouca
140	Media	Media	Muita	Hll	Muita	Pouca
141	Media	Muita	Pouca	Cir	Pouca	Muita
142	Media	Muita	Pouca	Cir	Muita	Muita
143	Media	Muita	Pouca	Cci	Pouca	Media-pouca
144	Media	Muita	Pouca	Cci	Muita	Media-pouca
145	Media	Muita	Pouca	Ali	Pouca	Media-pouca
146	Media	Muita	Pouca	Ali	Muita	Media-pouca
147	Media	Muita	Pouca	Ccc	Pouca	Media-pouca
148	Media	Muita	Pouca	Ccc	Muita	Media-pouca
149	Media	Muita	Pouca	Chb	Pouca	Media-pouca
150	Media	Muita	Pouca	Chb	Muita	Media-pouca
151	Media	Muita	Pouca	Chi	Pouca	Media-pouca
152	Media	Muita	Pouca	Chi	Muita	Media-pouca
153	Media	Muita	Pouca	Hll	Pouca	Media-pouca
154	Media	Muita	Pouca	Hll	Muita	Media-pouca
155	Media	Muita	Muita	Cir	Pouca	Muita
156	Media	Muita	Muita	Cir	Muita	Muita
157	Media	Muita	Muita	Cci	Pouca	Media-pouca
158	Media	Muita	Muita	Cci	Muita	Media-pouca
159	Media	Muita	Muita	Ali	Pouca	Media-pouca
160	Media	Muita	Muita	Ali	Muita	Media-pouca
161	Media	Muita	Muita	Ccc	Pouca	Media-pouca
162	Media	Muita	Muita	Ccc	Muita	Media-pouca

REGRA	CINZA	TAMANHO	ORIENTAÇÃO	FORMA	DISTÂNCIA	SALIÊNCIA TOPDOWN
163	Media	Muita	Muita	Chb	Pouca	Media-pouca
164	Media	Muita	Muita	Chb	Muita	Media-pouca
165	Media	Muita	Muita	Chi	Pouca	Media-pouca
166	Media	Muita	Muita	Chi	Muita	Media-pouca
167	Media	Muita	Muita	Hll	Pouca	Media-pouca
168	Media	Muita	Muita	Hll	Muita	Media-pouca
169	Muita	Pouca	Pouca	Cir	Pouca	Media
170	Muita	Pouca	Pouca	Cir	Muita	Media
171	Muita	Pouca	Pouca	Cci	Pouca	Pouca
172	Muita	Pouca	Pouca	Cci	Muita	Pouca
173	Muita	Pouca	Pouca	Ali	Pouca	Pouca
174	Muita	Pouca	Pouca	Ali	Muita	Pouca
175	Muita	Pouca	Pouca	Ccc	Pouca	Pouca
176	Muita	Pouca	Pouca	Ccc	Muita	Pouca
177	Muita	Pouca	Pouca	Chb	Pouca	Pouca
178	Muita	Pouca	Pouca	Chb	Muita	Pouca
179	Muita	Pouca	Pouca	Chi	Pouca	Pouca
180	Muita	Pouca	Pouca	Chi	Muita	Pouca
181	Muita	Pouca	Pouca	Hll	Pouca	Pouca
182	Muita	Pouca	Pouca	Hll	Muita	Pouca
183	Muita	Pouca	Muita	Cir	Pouca	Media
184	Muita	Pouca	Muita	Cir	Muita	Media
185	Muita	Pouca	Muita	Cci	Pouca	Pouca
186	Muita	Pouca	Muita	Cci	Muita	Pouca
187	Muita	Pouca	Muita	Ali	Pouca	Pouca
188	Muita	Pouca	Muita	Ali	Muita	Pouca
189	Muita	Pouca	Muita	Ccc	Pouca	Pouca
190	Muita	Pouca	Muita	Ccc	Muita	Pouca
191	Muita	Pouca	Muita	Chb	Pouca	Pouca
192	Muita	Pouca	Muita	Chb	Muita	Pouca
193	Muita	Pouca	Muita	Chi	Pouca	Pouca
194	Muita	Pouca	Muita	Chi	Muita	Pouca
195	Muita	Pouca	Muita	Hll	Pouca	Pouca
196	Muita	Pouca	Muita	Hll	Muita	Pouca
197	Muita	Média	Pouca	Cir	Pouca	Media-Muita
198	Muita	Media	Pouca	Cir	Muita	Média-Muita
199	Muita	Media	Pouca	Cci	Pouca	Pouca
200	Muita	Media	Pouca	Cci	Muita	Pouca
201	Muita	Media	Pouca	Ali	Pouca	Pouca
202	Muita	Media	Pouca	Ali	Muita	Pouca
203	Muita	Media	Pouca	Ccc	Pouca	Pouca
204	Muita	Media	Pouca	Ccc	Muita	Pouca
205	Muita	Media	Pouca	Chb	Pouca	Pouca
206	Muita	Media	Pouca	Chb	Muita	Pouca
207	Muita	Media	Pouca	Chi	Pouca	Pouca
208	Muita	Media	Pouca	Chi	Muita	Pouca
209	Muita	Media	Pouca	Hll	Pouca	Pouca
210	Muita	Media	Pouca	Hll	Muita	Pouca
211	Muita	Media	Muita	Cir	Pouca	Média-Muita
212	Muita	Media	Muita	Cir	Muita	Média-Muita

REGRA	CINZA	TAMANHO	ORIENTAÇÃO	FORMA	DISTÂNCIA	SALIÊNCIA TOPDOWN
213	Muita	Media	Muita	Cci	Pouca	Pouca
214	Muita	Media	Muita	Cci	Muita	Pouca
215	Muita	Media	Muita	Ali	Pouca	Pouca
216	Muita	Media	Muita	Ali	Muita	Pouca
217	Muita	Media	Muita	Ccc	Pouca	Pouca
218	Muita	Media	Muita	Ccc	Muita	Pouca
219	Muita	Media	Muita	Chb	Pouca	Pouca
220	Muita	Media	Muita	Chb	Muita	Pouca
221	Muita	Media	Muita	Chi	Pouca	Pouca
222	Muita	Media	Muita	Chi	Muita	Pouca
223	Muita	Media	Muita	Hll	Pouca	Pouca
224	Muita	Media	Muita	Hll	Muita	Pouca
225	Muita	Muita	Pouca	Cir	Pouca	Muita
226	Muita	Muita	Pouca	Cir	Muita	Muita
227	Muita	Muita	Pouca	Cci	Pouca	Media-pouca
228	Muita	Muita	Pouca	Cci	Muita	Media-pouca
229	Muita	Muita	Pouca	Ali	Pouca	Media-pouca
230	Muita	Muita	Pouca	Ali	Muita	Media-pouca
231	Muita	Muita	Pouca	Ccc	Pouca	Media-pouca
232	Muita	Muita	Pouca	Ccc	Muita	Media-pouca
233	Muita	Muita	Pouca	Chb	Pouca	Media-pouca
234	Muita	Muita	Pouca	Chb	Muita	Media-pouca
235	Muita	Muita	Pouca	Chi	Pouca	Media-pouca
236	Muita	Muita	Pouca	Chi	Muita	Media-pouca
237	Muita	Muita	Pouca	Hll	Pouca	Media-pouca
238	Muita	Muita	Pouca	Hll	Muita	Media-pouca
239	Muita	Muita	Muita	Cir	Pouca	Muita
240	Muita	Muita	Muita	Cir	Muita	Muita
241	Muita	Muita	Muita	Cci	Pouca	Media-pouca
242	Muita	Muita	Muita	Cci	Muita	Media-pouca
243	Muita	Muita	Muita	Ali	Pouca	Media-pouca
244	Muita	Muita	Muita	Ali	Muita	Media-pouca
245	Muita	Muita	Muita	Ccc	Pouca	Media-pouca
246	Muita	Muita	Muita	Ccc	Muita	Media-pouca
247	Muita	Muita	Muita	Chb	Pouca	Media-pouca
248	Muita	Muita	Muita	Chb	Muita	Media-pouca
249	Muita	Muita	Muita	Chi	Pouca	Media-pouca
250	Muita	Muita	Muita	Chi	Muita	Media-pouca
251	Muita	Muita	Muita	Hll	Pouca	Media-pouca
252	Muita	Muita	Muita	Hll	Muita	Media-pouca

Tabela A4.5 – Tabela de Regras para a tarefa de determinar os “Alicates Escuros”

REGRA	CINZA	TAMANHO	ORIENTAÇÃO	FORMA	DISTÂNCIA	SALIÊNCIA TOPDOWN
1	Pouca	Pouca	Pouca	Cir	Pouca	Media-pouca
2	Pouca	Pouca	Pouca	Cir	Muita	Media-pouca
3	Pouca	Pouca	Pouca	Cci	Pouca	Media-pouca
4	Pouca	Pouca	Pouca	Cci	Muita	Media-pouca
5	Pouca	Pouca	Pouca	Ali	Pouca	Muita
6	Pouca	Pouca	Pouca	Ali	Muita	Muita
7	Pouca	Pouca	Pouca	Ccc	Pouca	Media-pouca
8	Pouca	Pouca	Pouca	Ccc	Muita	Media-pouca
9	Pouca	Pouca	Pouca	Chb	Pouca	Media-pouca
10	Pouca	Pouca	Pouca	Chb	Muita	Media-pouca
11	Pouca	Pouca	Pouca	Chi	Pouca	Media-pouca
12	Pouca	Pouca	Pouca	Chi	Muita	Media-pouca
13	Pouca	Pouca	Pouca	Hll	Pouca	Media-pouca
14	Pouca	Pouca	Pouca	Hll	Muita	Media-pouca
15	Pouca	Pouca	Muita	Cir	Pouca	Media-pouca
16	Pouca	Pouca	Muita	Cir	Muita	Media-pouca
17	Pouca	Pouca	Muita	Cci	Pouca	Media-pouca
18	Pouca	Pouca	Muita	Cci	Muita	Media-pouca
19	Pouca	Pouca	Muita	Ali	Pouca	Muita
20	Pouca	Pouca	Muita	Ali	Muita	Muita
21	Pouca	Pouca	Muita	Ccc	Pouca	Media-pouca
22	Pouca	Pouca	Muita	Ccc	Muita	Media-pouca
23	Pouca	Pouca	Muita	Chb	Pouca	Media-pouca
24	Pouca	Pouca	Muita	Chb	Muita	Media-pouca
25	Pouca	Pouca	Muita	Chi	Pouca	Media-pouca
26	Pouca	Pouca	Muita	Chi	Muita	Media-pouca
27	Pouca	Pouca	Muita	Hll	Pouca	Media-pouca
28	Pouca	Pouca	Muita	Hll	Muita	Media-pouca
29	Pouca	Média	Pouca	Cir	Pouca	Media-pouca
30	Pouca	Media	Pouca	Cir	Muita	Media-pouca
31	Pouca	Media	Pouca	Cci	Pouca	Media-pouca
32	Pouca	Media	Pouca	Cci	Muita	Media-pouca
33	Pouca	Media	Pouca	Ali	Pouca	Muita
34	Pouca	Media	Pouca	Ali	Muita	Muita
35	Pouca	Media	Pouca	Ccc	Pouca	Media-pouca
36	Pouca	Media	Pouca	Ccc	Muita	Media-pouca
37	Pouca	Media	Pouca	Chb	Pouca	Media-pouca
38	Pouca	Media	Pouca	Chb	Muita	Media-pouca
39	Pouca	Media	Pouca	Chi	Pouca	Media-pouca
40	Pouca	Media	Pouca	Chi	Muita	Media-pouca
41	Pouca	Media	Pouca	Hll	Pouca	Media-pouca
42	Pouca	Media	Pouca	Hll	Muita	Media-pouca
43	Pouca	Media	Muita	Cir	Pouca	Media-pouca
44	Pouca	Media	Muita	Cir	Muita	Media-pouca
45	Pouca	Media	Muita	Cci	Pouca	Media-pouca
46	Pouca	Media	Muita	Cci	Muita	Media-pouca
47	Pouca	Media	Muita	Ali	Pouca	Muita
48	Pouca	Media	Muita	Ali	Muita	Muita

REGRA	CINZA	TAMANHO	ORIENTAÇÃO	FORMA	DISTÂNCIA	SALIÊNCIA TOPDOWN
49	Pouca	Media	Muita	Ccc	Pouca	Media-pouca
50	Pouca	Media	Muita	Ccc	Muita	Media-pouca
51	Pouca	Media	Muita	Chb	Pouca	Media-pouca
52	Pouca	Media	Muita	Chb	Muita	Media-pouca
53	Pouca	Media	Muita	Chi	Pouca	Media-pouca
54	Pouca	Media	Muita	Chi	Muita	Media-pouca
55	Pouca	Media	Muita	Hll	Pouca	Media-pouca
56	Pouca	Media	Muita	Hll	Muita	Media-pouca
57	Pouca	Muita	Pouca	Cir	Pouca	Media-pouca
58	Pouca	Muita	Pouca	Cir	Muita	Media-pouca
59	Pouca	Muita	Pouca	Cci	Pouca	Media-pouca
60	Pouca	Muita	Pouca	Cci	Muita	Media-pouca
61	Pouca	Muita	Pouca	Ali	Pouca	Muita
62	Pouca	Muita	Pouca	Ali	Muita	Muita
63	Pouca	Muita	Pouca	Ccc	Pouca	Media-pouca
64	Pouca	Muita	Pouca	Ccc	Muita	Media-pouca
65	Pouca	Muita	Pouca	Chb	Pouca	Media-pouca
66	Pouca	Muita	Pouca	Chb	Muita	Media-pouca
67	Pouca	Muita	Pouca	Chi	Pouca	Media-pouca
68	Pouca	Muita	Pouca	Chi	Muita	Media-pouca
69	Pouca	Muita	Pouca	Hll	Pouca	Media-pouca
70	Pouca	Muita	Pouca	Hll	Muita	Media-pouca
71	Pouca	Muita	Muita	Cir	Pouca	Media-pouca
72	Pouca	Muita	Muita	Cir	Muita	Media-pouca
73	Pouca	Muita	Muita	Cci	Pouca	Media-pouca
74	Pouca	Muita	Muita	Cci	Muita	Media-pouca
75	Pouca	Muita	Muita	Ali	Pouca	Muita
76	Pouca	Muita	Muita	Ali	Muita	Muita
77	Pouca	Muita	Muita	Ccc	Pouca	Media-pouca
78	Pouca	Muita	Muita	Ccc	Muita	Media-pouca
79	Pouca	Muita	Muita	Chb	Pouca	Media-pouca
80	Pouca	Muita	Muita	Chb	Muita	Media-pouca
81	Pouca	Muita	Muita	Chi	Pouca	Media-pouca
82	Pouca	Muita	Muita	Chi	Muita	Media-pouca
83	Pouca	Muita	Muita	Hll	Pouca	Media-pouca
84	Pouca	Muita	Muita	Hll	Muita	Media-pouca
85	Média	Pouca	Pouca	Cir	Pouca	Pouca
86	Media	Pouca	Pouca	Cir	Muita	Pouca
87	Media	Pouca	Pouca	Cci	Pouca	Pouca
88	Media	Pouca	Pouca	Cci	Muita	Pouca
89	Media	Pouca	Pouca	Ali	Pouca	Media-Muita
90	Media	Pouca	Pouca	Ali	Muita	Media-Muita
91	Media	Pouca	Pouca	Ccc	Pouca	Pouca
92	Media	Pouca	Pouca	Ccc	Muita	Pouca
93	Media	Pouca	Pouca	Chb	Pouca	Pouca
94	Media	Pouca	Pouca	Chb	Muita	Pouca
95	Media	Pouca	Pouca	Chi	Pouca	Pouca
96	Media	Pouca	Pouca	Chi	Muita	Pouca
97	Media	Pouca	Pouca	Hll	Pouca	Pouca
98	Media	Pouca	Pouca	Hll	Muita	Pouca

REGRA	CINZA	TAMANHO	ORIENTAÇÃO	FORMA	DISTÂNCIA	SALIÊNCIA TOPDOWN
99	Media	Pouca	Muita	Cir	Pouca	Pouca
100	Media	Pouca	Muita	Cir	Muita	Pouca
101	Media	Pouca	Muita	Cci	Pouca	Pouca
102	Media	Pouca	Muita	Cci	Muita	Pouca
103	Media	Pouca	Muita	Ali	Pouca	Media-Muita
104	Media	Pouca	Muita	Ali	Muita	Media-Muita
105	Media	Pouca	Muita	Ccc	Pouca	Pouca
106	Media	Pouca	Muita	Ccc	Muita	Pouca
107	Media	Pouca	Muita	Chb	Pouca	Pouca
108	Media	Pouca	Muita	Chb	Muita	Pouca
109	Media	Pouca	Muita	Chi	Pouca	Pouca
110	Media	Pouca	Muita	Chi	Muita	Pouca
111	Media	Pouca	Muita	Hll	Pouca	Pouca
112	Media	Pouca	Muita	Hll	Muita	Pouca
113	Média	Média	Pouca	Cir	Pouca	Pouca
114	Media	Media	Pouca	Cir	Muita	Pouca
115	Media	Media	Pouca	Cci	Pouca	Pouca
116	Media	Media	Pouca	Cci	Muita	Pouca
117	Media	Media	Pouca	Ali	Pouca	Media-Muita
118	Media	Media	Pouca	Ali	Muita	Media-Muita
119	Media	Media	Pouca	Ccc	Pouca	Pouca
120	Media	Media	Pouca	Ccc	Muita	Pouca
121	Media	Media	Pouca	Chb	Pouca	Pouca
122	Media	Media	Pouca	Chb	Muita	Pouca
123	Media	Media	Pouca	Chi	Pouca	Pouca
124	Media	Media	Pouca	Chi	Muita	Pouca
124	Media	Media	Pouca	Hll	Pouca	Pouca
126	Media	Media	Pouca	Hll	Muita	Pouca
127	Media	Media	Muita	Cir	Pouca	Pouca
128	Media	Media	Muita	Cir	Muita	Pouca
129	Media	Media	Muita	Cci	Pouca	Pouca
130	Media	Media	Muita	Cci	Muita	Pouca
131	Media	Media	Muita	Ali	Pouca	Media-Muita
132	Media	Media	Muita	Ali	Muita	Media-Muita
133	Media	Media	Muita	Ccc	Pouca	Pouca
134	Media	Media	Muita	Ccc	Muita	Pouca
135	Media	Media	Muita	Chb	Pouca	Pouca
136	Media	Media	Muita	Chb	Muita	Pouca
137	Media	Media	Muita	Chi	Pouca	Pouca
138	Media	Media	Muita	Chi	Muita	Pouca
139	Media	Media	Muita	Hll	Pouca	Pouca
140	Media	Media	Muita	Hll	Muita	Pouca
141	Media	Muita	Pouca	Cir	Pouca	Pouca
142	Media	Muita	Pouca	Cir	Muita	Pouca
143	Media	Muita	Pouca	Cci	Pouca	Pouca
144	Media	Muita	Pouca	Cci	Muita	Pouca
145	Media	Muita	Pouca	Ali	Pouca	Media-Muita
146	Media	Muita	Pouca	Ali	Muita	Media-Muita
147	Media	Muita	Pouca	Ccc	Pouca	Pouca
148	Media	Muita	Pouca	Ccc	Muita	Pouca

REGRA	CINZA	TAMANHO	ORIENTAÇÃO	FORMA	DISTÂNCIA	SALIÊNCIA TOPDOWN
149	Media	Muita	Pouca	Chb	Pouca	Pouca
150	Media	Muita	Pouca	Chb	Muita	Pouca
151	Media	Muita	Pouca	Chi	Pouca	Pouca
152	Media	Muita	Pouca	Chi	Muita	Pouca
153	Media	Muita	Pouca	Hll	Pouca	Pouca
154	Media	Muita	Pouca	Hll	Muita	Pouca
155	Media	Muita	Muita	Cir	Pouca	Pouca
156	Media	Muita	Muita	Cir	Muita	Pouca
157	Media	Muita	Muita	Cci	Pouca	Pouca
158	Media	Muita	Muita	Cci	Muita	Pouca
159	Media	Muita	Muita	Ali	Pouca	Media-Muita
160	Media	Muita	Muita	Ali	Muita	Media-Muita
161	Media	Muita	Muita	Ccc	Pouca	Pouca
162	Media	Muita	Muita	Ccc	Muita	Pouca
163	Media	Muita	Muita	Chb	Pouca	Pouca
164	Media	Muita	Muita	Chb	Muita	Pouca
165	Media	Muita	Muita	Chi	Pouca	Pouca
166	Media	Muita	Muita	Chi	Muita	Pouca
167	Media	Muita	Muita	Hll	Pouca	Pouca
168	Media	Muita	Muita	Hll	Muita	Pouca
169	Muita	Pouca	Pouca	Cir	Pouca	Pouca
170	Muita	Pouca	Pouca	Cir	Muita	Pouca
171	Muita	Pouca	Pouca	Cci	Pouca	Pouca
172	Muita	Pouca	Pouca	Cci	Muita	Pouca
173	Muita	Pouca	Pouca	Ali	Pouca	Media
174	Muita	Pouca	Pouca	Ali	Muita	Media
175	Muita	Pouca	Pouca	Ccc	Pouca	Pouca
176	Muita	Pouca	Pouca	Ccc	Muita	Pouca
177	Muita	Pouca	Pouca	Chb	Pouca	Pouca
178	Muita	Pouca	Pouca	Chb	Muita	Pouca
179	Muita	Pouca	Pouca	Chi	Pouca	Pouca
180	Muita	Pouca	Pouca	Chi	Muita	Pouca
181	Muita	Pouca	Pouca	Hll	Pouca	Pouca
182	Muita	Pouca	Pouca	Hll	Muita	Pouca
183	Muita	Pouca	Muita	Cir	Pouca	Pouca
184	Muita	Pouca	Muita	Cir	Muita	Pouca
185	Muita	Pouca	Muita	Cci	Pouca	Pouca
186	Muita	Pouca	Muita	Cci	Muita	Pouca
187	Muita	Pouca	Muita	Ali	Pouca	Media
188	Muita	Pouca	Muita	Ali	Muita	Media
189	Muita	Pouca	Muita	Ccc	Pouca	Pouca
190	Muita	Pouca	Muita	Ccc	Muita	Pouca
191	Muita	Pouca	Muita	Chb	Pouca	Pouca
192	Muita	Pouca	Muita	Chb	Muita	Pouca
193	Muita	Pouca	Muita	Chi	Pouca	Pouca
194	Muita	Pouca	Muita	Chi	Muita	Pouca
195	Muita	Pouca	Muita	Hll	Pouca	Pouca
196	Muita	Pouca	Muita	Hll	Muita	Pouca
197	Muita	Média	Pouca	Cir	Pouca	Pouca
198	Muita	Media	Pouca	Cir	Muita	Pouca

12/11/2019 14:00:00
12/11/2019 14:00:00

REGRA	CINZA	TAMANHO	ORIENTAÇÃO	FORMA	DISTÂNCIA	SALIÊNCIA TOPDOWN
199	Muita	Media	Pouca	Cci	Pouca	Pouca
200	Muita	Media	Pouca	Cci	Muita	Pouca
201	Muita	Media	Pouca	Ali	Pouca	Media
202	Muita	Media	Pouca	Ali	Muita	Media
203	Muita	Media	Pouca	Ccc	Pouca	Pouca
204	Muita	Media	Pouca	Ccc	Muita	Pouca
205	Muita	Media	Pouca	Chb	Pouca	Pouca
206	Muita	Media	Pouca	Chb	Muita	Pouca
207	Muita	Media	Pouca	Chi	Pouca	Pouca
208	Muita	Media	Pouca	Chi	Muita	Pouca
209	Muita	Media	Pouca	Hll	Pouca	Pouca
210	Muita	Media	Pouca	Hll	Muita	Pouca
211	Muita	Media	Muita	Cir	Pouca	Pouca
212	Muita	Media	Muita	Cir	Muita	Pouca
213	Muita	Media	Muita	Cci	Pouca	Pouca
214	Muita	Media	Muita	Cci	Muita	Pouca
215	Muita	Media	Muita	Ali	Pouca	Media
216	Muita	Media	Muita	Ali	Muita	Media
217	Muita	Media	Muita	Ccc	Pouca	Pouca
218	Muita	Media	Muita	Ccc	Muita	Pouca
219	Muita	Media	Muita	Chb	Pouca	Pouca
220	Muita	Media	Muita	Chb	Muita	Pouca
221	Muita	Media	Muita	Chi	Pouca	Pouca
222	Muita	Media	Muita	Chi	Muita	Pouca
223	Muita	Media	Muita	Hll	Pouca	Pouca
224	Muita	Media	Muita	Hll	Muita	Pouca
225	Muita	Muita	Pouca	Cir	Pouca	Pouca
226	Muita	Muita	Pouca	Cir	Muita	Pouca
227	Muita	Muita	Pouca	Cci	Pouca	Pouca
228	Muita	Muita	Pouca	Cci	Muita	Pouca
229	Muita	Muita	Pouca	Ali	Pouca	Media
230	Muita	Muita	Pouca	Ali	Muita	Media
231	Muita	Muita	Pouca	Ccc	Pouca	Pouca
232	Muita	Muita	Pouca	Ccc	Muita	Pouca
233	Muita	Muita	Pouca	Chb	Pouca	Pouca
234	Muita	Muita	Pouca	Chb	Muita	Pouca
235	Muita	Muita	Pouca	Chi	Pouca	Pouca
236	Muita	Muita	Pouca	Chi	Muita	Pouca
237	Muita	Muita	Pouca	Hll	Pouca	Pouca
238	Muita	Muita	Pouca	Hll	Muita	Pouca
239	Muita	Muita	Muita	Cir	Pouca	Pouca
240	Muita	Muita	Muita	Cir	Muita	Pouca
241	Muita	Muita	Muita	Cci	Pouca	Pouca
242	Muita	Muita	Muita	Cci	Muita	Pouca
243	Muita	Muita	Muita	Ali	Pouca	Media
244	Muita	Muita	Muita	Ali	Muita	Media
245	Muita	Muita	Muita	Ccc	Pouca	Pouca
246	Muita	Muita	Muita	Ccc	Muita	Pouca
247	Muita	Muita	Muita	Chb	Pouca	Pouca
248	Muita	Muita	Muita	Chb	Muita	Pouca

REGRA	CINZA	TAMANHO	ORIENTAÇÃO	FORMA	DISTÂNCIA	SALIÊNCIA TOPDOWN
249	Muita	Muita	Muita	Chi	Pouca	Pouca
250	Muita	Muita	Muita	Chi	Muita	Pouca
251	Muita	Muita	Muita	Hll	Pouca	Pouca
252	Muita	Muita	Muita	Hll	Muita	Pouca

Tabela A4.6 – Tabela de Regras para a tarefa de determinar o "Grifo (Chi) Escuro e Inclinado"

REGRA	CINZA	TAMANHO	ORIENTAÇÃO	FORMA	DISTÂNCIA	SALIÊNCIA TOPDOWN
1	Pouca	Pouca	Pouca	Cir	Pouca	Pouca
2	Pouca	Pouca	Pouca	Cir	Muita	Pouca
3	Pouca	Pouca	Pouca	Cci	Pouca	Pouca
4	Pouca	Pouca	Pouca	Cci	Muita	Pouca
5	Pouca	Pouca	Pouca	Ali	Pouca	Pouca
6	Pouca	Pouca	Pouca	Ali	Muita	Pouca
7	Pouca	Pouca	Pouca	Ccc	Pouca	Pouca
8	Pouca	Pouca	Pouca	Ccc	Muita	Pouca
9	Pouca	Pouca	Pouca	Chb	Pouca	Pouca
10	Pouca	Pouca	Pouca	Chb	Muita	Pouca
11	Pouca	Pouca	Pouca	Chi	Pouca	Media-Muita
12	Pouca	Pouca	Pouca	Chi	Muita	Média-Muita
13	Pouca	Pouca	Pouca	Hll	Pouca	Pouca
14	Pouca	Pouca	Pouca	Hll	Muita	Pouca
15	Pouca	Pouca	Muita	Cir	Pouca	Pouca
16	Pouca	Pouca	Muita	Cir	Muita	Pouca
17	Pouca	Pouca	Muita	Cci	Pouca	Pouca
18	Pouca	Pouca	Muita	Cci	Muita	Pouca
19	Pouca	Pouca	Muita	Ali	Pouca	Pouca
20	Pouca	Pouca	Muita	Ali	Muita	Pouca
21	Pouca	Pouca	Muita	Ccc	Pouca	Pouca
22	Pouca	Pouca	Muita	Ccc	Muita	Pouca
23	Pouca	Pouca	Muita	Chb	Pouca	Pouca
24	Pouca	Pouca	Muita	Chb	Muita	Pouca
25	Pouca	Pouca	Muita	Chi	Pouca	Muita
26	Pouca	Pouca	Muita	Chi	Muita	Muita
27	Pouca	Pouca	Muita	Hll	Pouca	Pouca
28	Pouca	Pouca	Muita	Hll	Muita	Pouca
29	Pouca	Média	Pouca	Cir	Pouca	Pouca
30	Pouca	Media	Pouca	Cir	Muita	Pouca
31	Pouca	Media	Pouca	Cci	Pouca	Pouca
32	Pouca	Media	Pouca	Cci	Muita	Pouca
33	Pouca	Media	Pouca	Ali	Pouca	Pouca
34	Pouca	Media	Pouca	Ali	Muita	Pouca
35	Pouca	Media	Pouca	Ccc	Pouca	Pouca
36	Pouca	Media	Pouca	Ccc	Muita	Pouca
37	Pouca	Media	Pouca	Chb	Pouca	Pouca
38	Pouca	Media	Pouca	Chb	Muita	Pouca
39	Pouca	Media	Pouca	Chi	Pouca	Media-Muita

REGRA	CINZA	TAMANHO	ORIENTAÇÃO	FORMA	DISTÂNCIA	SALIÊNCIA TOPDOWN
40	Pouca	Media	Pouca	Chi	Muita	Média-Muita
41	Pouca	Media	Pouca	Hll	Pouca	Pouca
42	Pouca	Media	Pouca	Hll	Muita	Pouca
43	Pouca	Media	Muita	Cir	Pouca	Pouca
44	Pouca	Media	Muita	Cir	Muita	Pouca
45	Pouca	Media	Muita	Cci	Pouca	Pouca
46	Pouca	Media	Muita	Cci	Muita	Pouca
47	Pouca	Media	Muita	Ali	Pouca	Pouca
48	Pouca	Media	Muita	Ali	Muita	Pouca
49	Pouca	Media	Muita	Ccc	Pouca	Pouca
50	Pouca	Media	Muita	Ccc	Muita	Pouca
51	Pouca	Media	Muita	Chb	Pouca	Pouca
52	Pouca	Media	Muita	Chb	Muita	Pouca
53	Pouca	Media	Muita	Chi	Pouca	Muita
54	Pouca	Media	Muita	Chi	Muita	Muita
55	Pouca	Media	Muita	Hll	Pouca	Pouca
56	Pouca	Media	Muita	Hll	Muita	Pouca
57	Pouca	Muita	Pouca	Cir	Pouca	Pouca
58	Pouca	Muita	Pouca	Cir	Muita	Pouca
59	Pouca	Muita	Pouca	Cci	Pouca	Pouca
60	Pouca	Muita	Pouca	Cci	Muita	Pouca
61	Pouca	Muita	Pouca	Ali	Pouca	Pouca
62	Pouca	Muita	Pouca	Ali	Muita	Pouca
63	Pouca	Muita	Pouca	Ccc	Pouca	Pouca
64	Pouca	Muita	Pouca	Ccc	Muita	Pouca
65	Pouca	Muita	Pouca	Chb	Pouca	Pouca
66	Pouca	Muita	Pouca	Chb	Muita	Pouca
67	Pouca	Muita	Pouca	Chi	Pouca	Media-Muita
68	Pouca	Muita	Pouca	Chi	Muita	Média-Muita
69	Pouca	Muita	Pouca	Hll	Pouca	Pouca
70	Pouca	Muita	Pouca	Hll	Muita	Pouca
71	Pouca	Muita	Muita	Cir	Pouca	Pouca
72	Pouca	Muita	Muita	Cir	Muita	Pouca
73	Pouca	Muita	Muita	Cci	Pouca	Pouca
74	Pouca	Muita	Muita	Cci	Muita	Pouca
75	Pouca	Muita	Muita	Ali	Pouca	Pouca
76	Pouca	Muita	Muita	Ali	Muita	Pouca
77	Pouca	Muita	Muita	Ccc	Pouca	Pouca
78	Pouca	Muita	Muita	Ccc	Muita	Pouca
79	Pouca	Muita	Muita	Chb	Pouca	Pouca
80	Pouca	Muita	Muita	Chb	Muita	Pouca
81	Pouca	Muita	Muita	Chi	Pouca	Muita
82	Pouca	Muita	Muita	Chi	Muita	Muita
83	Pouca	Muita	Muita	Hll	Pouca	Pouca
84	Pouca	Muita	Muita	Hll	Muita	Pouca
85	Média	Pouca	Pouca	Cir	Pouca	Pouca
86	Media	Pouca	Pouca	Cir	Muita	Pouca
87	Media	Pouca	Pouca	Cci	Pouca	Pouca
88	Media	Pouca	Pouca	Cci	Muita	Pouca
89	Media	Pouca	Pouca	Ali	Pouca	Pouca

REGRA	CINZA	TAMANHO	ORIENTAÇÃO	FORMA	DISTÂNCIA	SALIÊNCIA TOPDOWN
90	Media	Pouca	Pouca	Ali	Muita	Pouca
91	Media	Pouca	Pouca	Ccc	Pouca	Pouca
92	Media	Pouca	Pouca	Ccc	Muita	Pouca
93	Media	Pouca	Pouca	Chb	Pouca	Pouca
94	Media	Pouca	Pouca	Chb	Muita	Pouca
95	Media	Pouca	Pouca	Chi	Pouca	Media
96	Media	Pouca	Pouca	Chi	Muita	Media
97	Media	Pouca	Pouca	Hll	Pouca	Pouca
98	Media	Pouca	Pouca	Hll	Muita	Pouca
99	Media	Pouca	Muita	Cir	Pouca	Pouca
100	Media	Pouca	Muita	Cir	Muita	Pouca
101	Media	Pouca	Muita	Cci	Pouca	Pouca
102	Media	Pouca	Muita	Cci	Muita	Pouca
103	Media	Pouca	Muita	Ali	Pouca	Pouca
104	Media	Pouca	Muita	Ali	Muita	Pouca
105	Media	Pouca	Muita	Ccc	Pouca	Pouca
106	Media	Pouca	Muita	Ccc	Muita	Pouca
107	Media	Pouca	Muita	Chb	Pouca	Pouca
108	Media	Pouca	Muita	Chb	Muita	Pouca
109	Media	Pouca	Muita	Chi	Pouca	Media-Muita
110	Media	Pouca	Muita	Chi	Muita	Media-muita
111	Media	Pouca	Muita	Hll	Pouca	Pouca
112	Media	Pouca	Muita	Hll	Muita	Pouca
113	Média	Média	Pouca	Cir	Pouca	Pouca
114	Media	Media	Pouca	Cir	Muita	Pouca
115	Media	Media	Pouca	Cci	Pouca	Pouca
116	Media	Media	Pouca	Cci	Muita	Pouca
117	Media	Media	Pouca	Ali	Pouca	Pouca
118	Media	Media	Pouca	Ali	Muita	Pouca
119	Media	Media	Pouca	Ccc	Pouca	Pouca
120	Media	Media	Pouca	Ccc	Muita	Pouca
121	Media	Media	Pouca	Chb	Pouca	Pouca
122	Media	Media	Pouca	Chb	Muita	Pouca
123	Media	Media	Pouca	Chi	Pouca	Media
124	Media	Media	Pouca	Chi	Muita	Media
124	Media	Media	Pouca	Hll	Pouca	Pouca
126	Media	Media	Pouca	Hll	Muita	Pouca
127	Media	Media	Muita	Cir	Pouca	Pouca
128	Media	Media	Muita	Cir	Muita	Pouca
129	Media	Media	Muita	Cci	Pouca	Pouca
130	Media	Media	Muita	Cci	Muita	Pouca
131	Media	Media	Muita	Ali	Pouca	Pouca
132	Media	Media	Muita	Ali	Muita	Pouca
133	Media	Media	Muita	Ccc	Pouca	Pouca
134	Media	Media	Muita	Ccc	Muita	Pouca
135	Media	Media	Muita	Chb	Pouca	Pouca
136	Media	Media	Muita	Chb	Muita	Pouca
137	Media	Media	Muita	Chi	Pouca	Media-Muita
138	Media	Media	Muita	Chi	Muita	Média-Muita
139	Media	Media	Muita	Hll	Pouca	Pouca

REGRA	CINZA	TAMANHO	ORIENTAÇÃO	FORMA	DISTÂNCIA	SALIÊNCIA TOPDOWN
140	Media	Media	Muita	Hll	Muita	Pouca
141	Media	Muita	Pouca	Cir	Pouca	Pouca
142	Media	Muita	Pouca	Cir	Muita	Pouca
143	Media	Muita	Pouca	Cci	Pouca	Pouca
144	Media	Muita	Pouca	Cci	Muita	Pouca
145	Media	Muita	Pouca	Ali	Pouca	Pouca
146	Media	Muita	Pouca	Ali	Muita	Pouca
147	Media	Muita	Pouca	Ccc	Pouca	Pouca
148	Media	Muita	Pouca	Ccc	Muita	Pouca
149	Media	Muita	Pouca	Chb	Pouca	Pouca
150	Media	Muita	Pouca	Chb	Muita	Pouca
151	Media	Muita	Pouca	Chi	Pouca	Media
152	Media	Muita	Pouca	Chi	Muita	Média
153	Media	Muita	Pouca	Hll	Pouca	Pouca
154	Media	Muita	Pouca	Hll	Muita	Pouca
155	Media	Muita	Muita	Cir	Pouca	Pouca
156	Media	Muita	Muita	Cir	Muita	Pouca
157	Media	Muita	Muita	Cci	Pouca	Pouca
158	Media	Muita	Muita	Cci	Muita	Pouca
159	Media	Muita	Muita	Ali	Pouca	Pouca
160	Media	Muita	Muita	Ali	Muita	Pouca
161	Media	Muita	Muita	Ccc	Pouca	Pouca
162	Media	Muita	Muita	Ccc	Muita	Pouca
163	Media	Muita	Muita	Chb	Pouca	Pouca
164	Media	Muita	Muita	Chb	Muita	Pouca
165	Media	Muita	Muita	Chi	Pouca	Media-Muita
166	Media	Muita	Muita	Chi	Muita	Média-Muita
167	Media	Muita	Muita	Hll	Pouca	Pouca
168	Media	Muita	Muita	Hll	Muita	Pouca
169	Muita	Pouca	Pouca	Cir	Pouca	Pouca
170	Muita	Pouca	Pouca	Cir	Muita	Pouca
171	Muita	Pouca	Pouca	Cci	Pouca	Pouca
172	Muita	Pouca	Pouca	Cci	Muita	Pouca
173	Muita	Pouca	Pouca	Ali	Pouca	Pouca
174	Muita	Pouca	Pouca	Ali	Muita	Pouca
175	Muita	Pouca	Pouca	Ccc	Pouca	Pouca
176	Muita	Pouca	Pouca	Ccc	Muita	Pouca
177	Muita	Pouca	Pouca	Chb	Pouca	Pouca
178	Muita	Pouca	Pouca	Chb	Muita	Pouca
179	Muita	Pouca	Pouca	Chi	Pouca	Media-pouca
180	Muita	Pouca	Pouca	Chi	Muita	Média-pouca
181	Muita	Pouca	Pouca	Hll	Pouca	Pouca
182	Muita	Pouca	Pouca	Hll	Muita	Pouca
183	Muita	Pouca	Muita	Cir	Pouca	Pouca
184	Muita	Pouca	Muita	Cir	Muita	Pouca
185	Muita	Pouca	Muita	Cci	Pouca	Pouca
186	Muita	Pouca	Muita	Cci	Muita	Pouca
187	Muita	Pouca	Muita	Ali	Pouca	Pouca
188	Muita	Pouca	Muita	Ali	Muita	Pouca
189	Muita	Pouca	Muita	Ccc	Pouca	Pouca

REGRA	CINZA	TAMANHO	ORIENTAÇÃO	FORMA	DISTÂNCIA	SALIÊNCIA TOPDOWN
190	Muita	Pouca	Muita	Ccc	Muita	Pouca
191	Muita	Pouca	Muita	Chb	Pouca	Pouca
192	Muita	Pouca	Muita	Chb	Muita	Pouca
193	Muita	Pouca	Muita	Chi	Pouca	Média
194	Muita	Pouca	Muita	Chi	Muita	Média
195	Muita	Pouca	Muita	Hll	Pouca	Pouca
196	Muita	Pouca	Muita	Hll	Muita	Pouca
197	Muita	Média	Pouca	Cir	Pouca	Pouca
198	Muita	Media	Pouca	Cir	Muita	Pouca
199	Muita	Media	Pouca	Cci	Pouca	Pouca
200	Muita	Media	Pouca	Cci	Muita	Pouca
201	Muita	Media	Pouca	Ali	Pouca	Pouca
202	Muita	Media	Pouca	Ali	Muita	Pouca
203	Muita	Media	Pouca	Ccc	Pouca	Pouca
204	Muita	Media	Pouca	Ccc	Muita	Pouca
205	Muita	Media	Pouca	Chb	Pouca	Pouca
206	Muita	Media	Pouca	Chb	Muita	Pouca
207	Muita	Media	Pouca	Chi	Pouca	Media-pouca
208	Muita	Media	Pouca	Chi	Muita	Média-pouca
209	Muita	Media	Pouca	Hll	Pouca	Pouca
210	Muita	Media	Pouca	Hll	Muita	Pouca
211	Muita	Media	Muita	Cir	Pouca	Pouca
212	Muita	Media	Muita	Cir	Muita	Pouca
213	Muita	Media	Muita	Cci	Pouca	Pouca
214	Muita	Media	Muita	Cci	Muita	Pouca
215	Muita	Media	Muita	Ali	Pouca	Pouca
216	Muita	Media	Muita	Ali	Muita	Pouca
217	Muita	Media	Muita	Ccc	Pouca	Pouca
218	Muita	Media	Muita	Ccc	Muita	Pouca
219	Muita	Media	Muita	Chb	Pouca	Pouca
220	Muita	Media	Muita	Chb	Muita	Pouca
221	Muita	Media	Muita	Chi	Pouca	Média
222	Muita	Media	Muita	Chi	Muita	Média
223	Muita	Media	Muita	Hll	Pouca	Pouca
224	Muita	Media	Muita	Hll	Muita	Pouca
225	Muita	Muita	Pouca	Cir	Pouca	Pouca
226	Muita	Muita	Pouca	Cir	Muita	Pouca
227	Muita	Muita	Pouca	Cci	Pouca	Pouca
228	Muita	Muita	Pouca	Cci	Muita	Pouca
229	Muita	Muita	Pouca	Ali	Pouca	Pouca
230	Muita	Muita	Pouca	Ali	Muita	Pouca
231	Muita	Muita	Pouca	Ccc	Pouca	Pouca
232	Muita	Muita	Pouca	Ccc	Muita	Pouca
233	Muita	Muita	Pouca	Chb	Pouca	Pouca
234	Muita	Muita	Pouca	Chb	Muita	Pouca
235	Muita	Muita	Pouca	Chi	Pouca	Media-pouca
236	Muita	Muita	Pouca	Chi	Muita	Média-pouca
237	Muita	Muita	Pouca	Hll	Pouca	Pouca
238	Muita	Muita	Pouca	Hll	Muita	Pouca
239	Muita	Muita	Muita	Cir	Pouca	Pouca

REGRA	CINZA	TAMANHO	ORIENTAÇÃO	FORMA	DISTÂNCIA	SALIÊNCIA TOPDOWN
240	Muita	Muita	Muita	Cir	Muita	Pouca
241	Muita	Muita	Muita	Cci	Pouca	Pouca
242	Muita	Muita	Muita	Cci	Muita	Pouca
243	Muita	Muita	Muita	Ali	Pouca	Pouca
244	Muita	Muita	Muita	Ali	Muita	Pouca
245	Muita	Muita	Muita	Ccc	Pouca	Pouca
246	Muita	Muita	Muita	Ccc	Muita	Pouca
247	Muita	Muita	Muita	Chb	Pouca	Pouca
248	Muita	Muita	Muita	Chb	Muita	Pouca
249	Muita	Muita	Muita	Chi	Pouca	Media
250	Muita	Muita	Muita	Chi	Muita	Média
251	Muita	Muita	Muita	Hll	Pouca	Pouca
252	Muita	Muita	Muita	Hll	Muita	Pouca

Tabela A4.7 – Tabela de Regras para a tarefa de determinar a “Chava Inglesa (Chb) Clara, Pequena e Distante”

REGRA	CINZA	TAMANHO	ORIENTAÇÃO	FORMA	DISTANCIA	SALIENCIA TOPDOWN
1	Pouca	Pouca	Pouca	Cir	Pouca	Pouca
2	Pouca	Pouca	Pouca	Cir	Muita	Pouca
3	Pouca	Pouca	Pouca	Cci	Pouca	Pouca
4	Pouca	Pouca	Pouca	Cci	Muita	Pouca
5	Pouca	Pouca	Pouca	Ali	Pouca	Pouca
6	Pouca	Pouca	Pouca	Ali	Muita	Pouca
7	Pouca	Pouca	Pouca	Ccc	Pouca	Pouca
8	Pouca	Pouca	Pouca	Ccc	Muita	Pouca
9	Pouca	Pouca	Pouca	Chb	Pouca	Media-pouca
10	Pouca	Pouca	Pouca	Chb	Muita	Muita
11	Pouca	Pouca	Pouca	Chi	Pouca	Pouca
12	Pouca	Pouca	Pouca	Chi	Muita	Pouca
13	Pouca	Pouca	Pouca	Hll	Pouca	Pouca
14	Pouca	Pouca	Pouca	Hll	Muita	Pouca
15	Pouca	Pouca	Muita	Cir	Pouca	Pouca
16	Pouca	Pouca	Muita	Cir	Muita	Pouca
17	Pouca	Pouca	Muita	Cci	Pouca	Pouca
18	Pouca	Pouca	Muita	Cci	Muita	Pouca
19	Pouca	Pouca	Muita	Ali	Pouca	Pouca
20	Pouca	Pouca	Muita	Ali	Muita	Pouca
21	Pouca	Pouca	Muita	Ccc	Pouca	Pouca
22	Pouca	Pouca	Muita	Ccc	Muita	Pouca
23	Pouca	Pouca	Muita	Chb	Pouca	Media-pouca
24	Pouca	Pouca	Muita	Chb	Muita	Muita
25	Pouca	Pouca	Muita	Chi	Pouca	Pouca
26	Pouca	Pouca	Muita	Chi	Muita	Pouca
27	Pouca	Pouca	Muita	Hll	Pouca	Pouca
28	Pouca	Pouca	Muita	Hll	Muita	Pouca
29	Pouca	Média	Pouca	Cir	Pouca	Pouca
30	Pouca	Media	Pouca	Cir	Muita	Pouca

REGRA	CINZA	TAMANHO	ORIENTAÇÃO	FORMA	DISTANCIA	SALIENCIA TOPDOWN
31	Pouca	Media	Pouca	Cci	Pouca	Pouca
32	Pouca	Media	Pouca	Cci	Muita	Pouca
33	Pouca	Media	Pouca	Ali	Pouca	Pouca
34	Pouca	Media	Pouca	Ali	Muita	Pouca
35	Pouca	Media	Pouca	Ccc	Pouca	Pouca
36	Pouca	Media	Pouca	Ccc	Muita	Pouca
37	Pouca	Media	Pouca	Chb	Pouca	Media-pouca
38	Pouca	Media	Pouca	Chb	Muita	Media-pouca
39	Pouca	Media	Pouca	Chi	Pouca	Pouca
40	Pouca	Media	Pouca	Chi	Muita	Pouca
41	Pouca	Media	Pouca	Hll	Pouca	Pouca
42	Pouca	Media	Pouca	Hll	Muita	Pouca
43	Pouca	Media	Muita	Cir	Pouca	Pouca
44	Pouca	Media	Muita	Cir	Muita	Pouca
45	Pouca	Media	Muita	Cci	Pouca	Pouca
46	Pouca	Media	Muita	Cci	Muita	Pouca
47	Pouca	Media	Muita	Ali	Pouca	Pouca
48	Pouca	Media	Muita	Ali	Muita	Pouca
49	Pouca	Media	Muita	Ccc	Pouca	Pouca
50	Pouca	Media	Muita	Ccc	Muita	Pouca
51	Pouca	Media	Muita	Chb	Pouca	Media-pouca
52	Pouca	Media	Muita	Chb	Muita	Media-pouca
53	Pouca	Media	Muita	Chi	Pouca	Pouca
54	Pouca	Media	Muita	Chi	Muita	Pouca
55	Pouca	Media	Muita	Hll	Pouca	Pouca
56	Pouca	Media	Muita	Hll	Muita	Pouca
57	Pouca	Muita	Pouca	Cir	Pouca	Pouca
58	Pouca	Muita	Pouca	Cir	Muita	Pouca
59	Pouca	Muita	Pouca	Cci	Pouca	Pouca
60	Pouca	Muita	Pouca	Cci	Muita	Pouca
61	Pouca	Muita	Pouca	Ali	Pouca	Pouca
62	Pouca	Muita	Pouca	Ali	Muita	Pouca
63	Pouca	Muita	Pouca	Ccc	Pouca	Pouca
64	Pouca	Muita	Pouca	Ccc	Muita	Pouca
65	Pouca	Muita	Pouca	Chb	Pouca	Media-pouca
66	Pouca	Muita	Pouca	Chb	Muita	Media-pouca
67	Pouca	Muita	Pouca	Chi	Pouca	Pouca
68	Pouca	Muita	Pouca	Chi	Muita	Pouca
69	Pouca	Muita	Pouca	Hll	Pouca	Pouca
70	Pouca	Muita	Pouca	Hll	Muita	Pouca
71	Pouca	Muita	Muita	Cir	Pouca	Pouca
72	Pouca	Muita	Muita	Cir	Muita	Pouca
73	Pouca	Muita	Muita	Cci	Pouca	Pouca
74	Pouca	Muita	Muita	Cci	Muita	Pouca
75	Pouca	Muita	Muita	Ali	Pouca	Pouca
76	Pouca	Muita	Muita	Ali	Muita	Pouca
77	Pouca	Muita	Muita	Ccc	Pouca	Pouca
78	Pouca	Muita	Muita	Ccc	Muita	Pouca
79	Pouca	Muita	Muita	Chb	Pouca	Media-pouca
80	Pouca	Muita	Muita	Chb	Muita	Media-pouca

REGRA	CINZA	TAMANHO	ORIENTAÇÃO	FORMA	DISTANCIA	SALIENCIA TOPDOWN
81	Pouca	Muita	Muita	Chi	Pouca	Pouca
82	Pouca	Muita	Muita	Chi	Muita	Pouca
83	Pouca	Muita	Muita	Hll	Pouca	Pouca
84	Pouca	Muita	Muita	Hll	Muita	Pouca
85	Média	Pouca	Pouca	Cir	Pouca	Pouca
86	Media	Pouca	Pouca	Cir	Muita	Pouca
87	Media	Pouca	Pouca	Cci	Pouca	Pouca
88	Media	Pouca	Pouca	Cci	Muita	Pouca
89	Media	Pouca	Pouca	Ali	Pouca	Pouca
90	Media	Pouca	Pouca	Ali	Muita	Pouca
91	Media	Pouca	Pouca	Ccc	Pouca	Pouca
92	Media	Pouca	Pouca	Ccc	Muita	Pouca
93	Media	Pouca	Pouca	Chb	Pouca	Media
94	Media	Pouca	Pouca	Chb	Muita	Media-muita
95	Media	Pouca	Pouca	Chi	Pouca	Pouca
96	Media	Pouca	Pouca	Chi	Muita	Pouca
97	Media	Pouca	Pouca	Hll	Pouca	Pouca
98	Media	Pouca	Pouca	Hll	Muita	Pouca
99	Media	Pouca	Muita	Cir	Pouca	Pouca
100	Media	Pouca	Muita	Cir	Muita	Pouca
101	Media	Pouca	Muita	Cci	Pouca	Pouca
102	Media	Pouca	Muita	Cci	Muita	Pouca
103	Media	Pouca	Muita	Ali	Pouca	Pouca
104	Media	Pouca	Muita	Ali	Muita	Pouca
105	Media	Pouca	Muita	Ccc	Pouca	Pouca
106	Media	Pouca	Muita	Ccc	Muita	Pouca
107	Media	Pouca	Muita	Chb	Pouca	Media
108	Media	Pouca	Muita	Chb	Muita	Media-muita
109	Media	Pouca	Muita	Chi	Pouca	Pouca
110	Media	Pouca	Muita	Chi	Muita	Pouca
111	Media	Pouca	Muita	Hll	Pouca	Pouca
112	Media	Pouca	Muita	Hll	Muita	Pouca
113	Média	Média	Pouca	Cir	Pouca	Pouca
114	Media	Media	Pouca	Cir	Muita	Pouca
115	Media	Media	Pouca	Cci	Pouca	Pouca
116	Media	Media	Pouca	Cci	Muita	Pouca
117	Media	Media	Pouca	Ali	Pouca	Pouca
118	Media	Media	Pouca	Ali	Muita	Pouca
119	Media	Media	Pouca	Ccc	Pouca	Pouca
120	Media	Media	Pouca	Ccc	Muita	Pouca
121	Media	Media	Pouca	Chb	Pouca	Media-pouca
122	Media	Media	Pouca	Chb	Muita	Media
123	Media	Media	Pouca	Chi	Pouca	Pouca
124	Media	Media	Pouca	Chi	Muita	Pouca
124	Media	Media	Pouca	Hll	Pouca	Pouca
126	Media	Media	Pouca	Hll	Muita	Pouca
127	Media	Media	Muita	Cir	Pouca	Pouca
128	Media	Media	Muita	Cir	Muita	Pouca
129	Media	Media	Muita	Cci	Pouca	Pouca
130	Media	Media	Muita	Cci	Muita	Pouca

REGRA	CINZA	TAMANHO	ORIENTAÇÃO	FORMA	DISTANCIA	SALIENCIA TOPDOWN
131	Media	Media	Muita	Ali	Pouca	Pouca
132	Media	Media	Muita	Ali	Muita	Pouca
133	Media	Media	Muita	Ccc	Pouca	Pouca
134	Media	Media	Muita	Ccc	Muita	Pouca
135	Media	Media	Muita	Chb	Pouca	Media-pouca
136	Media	Media	Muita	Chb	Muita	Media
137	Media	Media	Muita	Chi	Pouca	Pouca
138	Media	Media	Muita	Chi	Muita	Pouca
139	Media	Media	Muita	Hll	Pouca	Pouca
140	Media	Media	Muita	Hll	Muita	Pouca
141	Media	Muita	Pouca	Cir	Pouca	Pouca
142	Media	Muita	Pouca	Cir	Muita	Pouca
143	Media	Muita	Pouca	Cci	Pouca	Pouca
144	Media	Muita	Pouca	Cci	Muita	Pouca
145	Media	Muita	Pouca	Ali	Pouca	Pouca
146	Media	Muita	Pouca	Ali	Muita	Pouca
147	Media	Muita	Pouca	Ccc	Pouca	Pouca
148	Media	Muita	Pouca	Ccc	Muita	Pouca
149	Media	Muita	Pouca	Chb	Pouca	Media-pouca
150	Media	Muita	Pouca	Chb	Muita	Media-pouca
151	Media	Muita	Pouca	Chi	Pouca	Pouca
152	Media	Muita	Pouca	Chi	Muita	Pouca
153	Media	Muita	Pouca	Hll	Pouca	Pouca
154	Media	Muita	Pouca	Hll	Muita	Pouca
155	Media	Muita	Muita	Cir	Pouca	Pouca
156	Media	Muita	Muita	Cir	Muita	Pouca
157	Media	Muita	Muita	Cci	Pouca	Pouca
158	Media	Muita	Muita	Cci	Muita	Pouca
159	Media	Muita	Muita	Ali	Pouca	Pouca
160	Media	Muita	Muita	Ali	Muita	Pouca
161	Media	Muita	Muita	Ccc	Pouca	Pouca
162	Media	Muita	Muita	Ccc	Muita	Pouca
163	Media	Muita	Muita	Chb	Pouca	Media-pouca
164	Media	Muita	Muita	Chb	Muita	Media-pouca
165	Media	Muita	Muita	Chi	Pouca	Pouca
166	Media	Muita	Muita	Chi	Muita	Pouca
167	Media	Muita	Muita	Hll	Pouca	Pouca
168	Media	Muita	Muita	Hll	Muita	Pouca
169	Muita	Pouca	Pouca	Cir	Pouca	Pouca
170	Muita	Pouca	Pouca	Cir	Muita	Pouca
171	Muita	Pouca	Pouca	Cci	Pouca	Pouca
172	Muita	Pouca	Pouca	Cci	Muita	Pouca
173	Muita	Pouca	Pouca	Ali	Pouca	Pouca
174	Muita	Pouca	Pouca	Ali	Muita	Pouca
175	Muita	Pouca	Pouca	Ccc	Pouca	Pouca
176	Muita	Pouca	Pouca	Ccc	Muita	Pouca
177	Muita	Pouca	Pouca	Chb	Pouca	Media-muita
178	Muita	Pouca	Pouca	Chb	Muita	Muita
179	Muita	Pouca	Pouca	Chi	Pouca	Pouca
180	Muita	Pouca	Pouca	Chi	Muita	Pouca

REGRA	CINZA	TAMANHO	ORIENTAÇÃO	FORMA	DISTANCIA	SALIENCIA TOPDOWN
181	Muita	Pouca	Pouca	Hll	Pouca	Pouca
182	Muita	Pouca	Pouca	Hll	Muita	Pouca
183	Muita	Pouca	Muita	Cir	Pouca	Pouca
184	Muita	Pouca	Muita	Cir	Muita	Pouca
185	Muita	Pouca	Muita	Cci	Pouca	Pouca
186	Muita	Pouca	Muita	Cci	Muita	Pouca
187	Muita	Pouca	Muita	Ali	Pouca	Pouca
188	Muita	Pouca	Muita	Ali	Muita	Pouca
189	Muita	Pouca	Muita	Ccc	Pouca	Pouca
190	Muita	Pouca	Muita	Ccc	Muita	Pouca
191	Muita	Pouca	Muita	Chb	Pouca	Media-muita
192	Muita	Pouca	Muita	Chb	Muita	Muita
193	Muita	Pouca	Muita	Chi	Pouca	Pouca
194	Muita	Pouca	Muita	Chi	Muita	Pouca
195	Muita	Pouca	Muita	Hll	Pouca	Pouca
196	Muita	Pouca	Muita	Hll	Muita	Pouca
197	Muita	Média	Pouca	Cir	Pouca	Pouca
198	Muita	Media	Pouca	Cir	Muita	Pouca
199	Muita	Media	Pouca	Cci	Pouca	Pouca
200	Muita	Media	Pouca	Cci	Muita	Pouca
201	Muita	Media	Pouca	Ali	Pouca	Pouca
202	Muita	Media	Pouca	Ali	Muita	Pouca
203	Muita	Media	Pouca	Ccc	Pouca	Pouca
204	Muita	Media	Pouca	Ccc	Muita	Pouca
205	Muita	Media	Pouca	Chb	Pouca	Media
206	Muita	Media	Pouca	Chb	Muita	Media-muita
207	Muita	Media	Pouca	Chi	Pouca	Pouca
208	Muita	Media	Pouca	Chi	Muita	Pouca
209	Muita	Media	Pouca	Hll	Pouca	Pouca
210	Muita	Media	Pouca	Hll	Muita	Pouca
211	Muita	Media	Muita	Cir	Pouca	Pouca
212	Muita	Media	Muita	Cir	Muita	Pouca
213	Muita	Media	Muita	Cci	Pouca	Pouca
214	Muita	Media	Muita	Cci	Muita	Pouca
215	Muita	Media	Muita	Ali	Pouca	Pouca
216	Muita	Media	Muita	Ali	Muita	Pouca
217	Muita	Media	Muita	Ccc	Pouca	Pouca
218	Muita	Media	Muita	Ccc	Muita	Pouca
219	Muita	Media	Muita	Chb	Pouca	Media
220	Muita	Media	Muita	Chb	Muita	Media-muita
221	Muita	Media	Muita	Chi	Pouca	Pouca
222	Muita	Media	Muita	Chi	Muita	Pouca
223	Muita	Media	Muita	Hll	Pouca	Pouca
224	Muita	Media	Muita	Hll	Muita	Pouca
225	Muita	Muita	Pouca	Cir	Pouca	Pouca
226	Muita	Muita	Pouca	Cir	Muita	Pouca
227	Muita	Muita	Pouca	Cci	Pouca	Pouca
228	Muita	Muita	Pouca	Cci	Muita	Pouca
229	Muita	Muita	Pouca	Ali	Pouca	Pouca
230	Muita	Muita	Pouca	Ali	Muita	Pouca

REGRA	CINZA	TAMANHO	ORIENTAÇÃO	FORMA	DISTANCIA	SALIENCIA TOPDOWN
231	Muita	Muita	Pouca	Ccc	Pouca	Pouca
232	Muita	Muita	Pouca	Ccc	Muita	Pouca
233	Muita	Muita	Pouca	Chb	Pouca	Media-pouca
234	Muita	Muita	Pouca	Chb	Muita	Media
235	Muita	Muita	Pouca	Chi	Pouca	Pouca
236	Muita	Muita	Pouca	Chi	Muita	Pouca
237	Muita	Muita	Pouca	Hll	Pouca	Pouca
238	Muita	Muita	Pouca	Hll	Muita	Pouca
239	Muita	Muita	Muita	Cir	Pouca	Pouca
240	Muita	Muita	Muita	Cir	Muita	Pouca
241	Muita	Muita	Muita	Cci	Pouca	Pouca
242	Muita	Muita	Muita	Cci	Muita	Pouca
243	Muita	Muita	Muita	Ali	Pouca	Pouca
244	Muita	Muita	Muita	Ali	Muita	Pouca
245	Muita	Muita	Muita	Ccc	Pouca	Pouca
246	Muita	Muita	Muita	Ccc	Muita	Pouca
247	Muita	Muita	Muita	Chb	Pouca	Media-pouca
248	Muita	Muita	Muita	Chb	Muita	Media
249	Muita	Muita	Muita	Chi	Pouca	Pouca
250	Muita	Muita	Muita	Chi	Muita	Pouca
251	Muita	Muita	Muita	Hll	Pouca	Pouca
252	Muita	Muita	Muita	Hll	Muita	Pouca

Tabela A4.8– Tabela de Regras para a tarefa de determinar a “Chave Hallen Escura, Grande, Inclinada e Próxima ”

REGRA	CINZA	TAMANHO	ORIENTAÇÃO	FORMA	DISTÂNCIA	SALIÊNCIA TOPDOWN
1	Pouca	Pouca	Pouca	Cir	Pouca	Pouca
2	Pouca	Pouca	Pouca	Cir	Muita	Pouca
3	Pouca	Pouca	Pouca	Cci	Pouca	Pouca
4	Pouca	Pouca	Pouca	Cci	Muita	Pouca
5	Pouca	Pouca	Pouca	Ali	Pouca	Pouca
6	Pouca	Pouca	Pouca	Ali	Muita	Pouca
7	Pouca	Pouca	Pouca	Ccc	Pouca	Pouca
8	Pouca	Pouca	Pouca	Ccc	Muita	Pouca
9	Pouca	Pouca	Pouca	Chb	Pouca	Pouca
10	Pouca	Pouca	Pouca	Chb	Muita	Pouca
11	Pouca	Pouca	Pouca	Chi	Pouca	Pouca
12	Pouca	Pouca	Pouca	Chi	Muita	Pouca
13	Pouca	Pouca	Pouca	Hll	Pouca	Media-pouca
14	Pouca	Pouca	Pouca	Hll	Muita	Media
15	Pouca	Pouca	Muita	Cir	Pouca	Pouca
16	Pouca	Pouca	Muita	Cir	Muita	Pouca
17	Pouca	Pouca	Muita	Cci	Pouca	Pouca
18	Pouca	Pouca	Muita	Cci	Muita	Pouca
19	Pouca	Pouca	Muita	Ali	Pouca	Pouca
20	Pouca	Pouca	Muita	Ali	Muita	Pouca
21	Pouca	Pouca	Muita	Ccc	Pouca	Pouca
22	Pouca	Pouca	Muita	Ccc	Muita	Pouca
23	Pouca	Pouca	Muita	Chb	Pouca	Pouca

REGRA	CINZA	TAMANHO	ORIENTAÇÃO	FORMA	DISTÂNCIA	SALIÊNCIA TOPDOWN
24	Pouca	Pouca	Muita	Chb	Muita	Pouca
25	Pouca	Pouca	Muita	Chi	Pouca	Pouca
26	Pouca	Pouca	Muita	Chi	Muita	Pouca
27	Pouca	Pouca	Muita	Hll	Pouca	Media
28	Pouca	Pouca	Muita	Hll	Muita	Media-muita
29	Pouca	Média	Pouca	Cir	Pouca	Pouca
30	Pouca	Media	Pouca	Cir	Muita	Pouca
31	Pouca	Media	Pouca	Cci	Pouca	Pouca
32	Pouca	Media	Pouca	Cci	Muita	Pouca
33	Pouca	Media	Pouca	Ali	Pouca	Pouca
34	Pouca	Media	Pouca	Ali	Muita	Pouca
35	Pouca	Media	Pouca	Ccc	Pouca	Pouca
36	Pouca	Media	Pouca	Ccc	Muita	Pouca
37	Pouca	Media	Pouca	Chb	Pouca	Pouca
38	Pouca	Media	Pouca	Chb	Muita	Pouca
39	Pouca	Media	Pouca	Chi	Pouca	Pouca
40	Pouca	Media	Pouca	Chi	Muita	Pouca
41	Pouca	Media	Pouca	Hll	Pouca	Media-pouca
42	Pouca	Media	Pouca	Hll	Muita	Media
43	Pouca	Media	Muita	Cir	Pouca	Pouca
44	Pouca	Media	Muita	Cir	Muita	Pouca
45	Pouca	Media	Muita	Cci	Pouca	Pouca
46	Pouca	Media	Muita	Cci	Muita	Pouca
47	Pouca	Media	Muita	Ali	Pouca	Pouca
48	Pouca	Media	Muita	Ali	Muita	Pouca
49	Pouca	Media	Muita	Ccc	Pouca	Pouca
50	Pouca	Media	Muita	Ccc	Muita	Pouca
51	Pouca	Media	Muita	Chb	Pouca	Pouca
52	Pouca	Media	Muita	Chb	Muita	Pouca
53	Pouca	Media	Muita	Chi	Pouca	Pouca
54	Pouca	Media	Muita	Chi	Muita	Pouca
55	Pouca	Media	Muita	Hll	Pouca	Media
56	Pouca	Media	Muita	Hll	Muita	Media-muita
57	Pouca	Muita	Pouca	Cir	Pouca	Pouca
58	Pouca	Muita	Pouca	Cir	Muita	Pouca
59	Pouca	Muita	Pouca	Cci	Pouca	Pouca
60	Pouca	Muita	Pouca	Cci	Muita	Pouca
61	Pouca	Muita	Pouca	Ali	Pouca	Pouca
62	Pouca	Muita	Pouca	Ali	Muita	Pouca
63	Pouca	Muita	Pouca	Ccc	Pouca	Pouca
64	Pouca	Muita	Pouca	Ccc	Muita	Pouca
65	Pouca	Muita	Pouca	Chb	Pouca	Pouca
66	Pouca	Muita	Pouca	Chb	Muita	Pouca
67	Pouca	Muita	Pouca	Chi	Pouca	Pouca
68	Pouca	Muita	Pouca	Chi	Muita	Pouca
69	Pouca	Muita	Pouca	Hll	Pouca	Media
70	Pouca	Muita	Pouca	Hll	Muita	Media-muita
71	Pouca	Muita	Muita	Cir	Pouca	Pouca
72	Pouca	Muita	Muita	Cir	Muita	Pouca
73	Pouca	Muita	Muita	Cci	Pouca	Pouca
74	Pouca	Muita	Muita	Cci	Muita	Pouca
75	Pouca	Muita	Muita	Ali	Pouca	Pouca
76	Pouca	Muita	Muita	Ali	Muita	Pouca
77	Pouca	Muita	Muita	Ccc	Pouca	Pouca
78	Pouca	Muita	Muita	Ccc	Muita	Pouca

REGRA	CINZA	TAMANHO	ORIENTAÇÃO	FORMA	DISTÂNCIA	SALIÊNCIA TOPDOWN
79	Pouca	Muita	Muita	Chb	Pouca	Pouca
80	Pouca	Muita	Muita	Chb	Muita	Pouca
81	Pouca	Muita	Muita	Chi	Pouca	Pouca
82	Pouca	Muita	Muita	Chi	Muita	Pouca
83	Pouca	Muita	Muita	Hll	Pouca	Muita
84	Pouca	Muita	Muita	Hll	Muita	Media-muita
85	Média	Pouca	Pouca	Cir	Pouca	Pouca
86	Media	Pouca	Pouca	Cir	Muita	Pouca
87	Media	Pouca	Pouca	Cci	Pouca	Pouca
88	Media	Pouca	Pouca	Cci	Muita	Pouca
89	Media	Pouca	Pouca	Ali	Pouca	Pouca
90	Media	Pouca	Pouca	Ali	Muita	Pouca
91	Media	Pouca	Pouca	Ccc	Pouca	Pouca
92	Media	Pouca	Pouca	Ccc	Muita	Pouca
93	Media	Pouca	Pouca	Chb	Pouca	Pouca
94	Media	Pouca	Pouca	Chb	Muita	Pouca
95	Media	Pouca	Pouca	Chi	Pouca	Pouca
96	Media	Pouca	Pouca	Chi	Muita	Pouca
97	Media	Pouca	Pouca	Hll	Pouca	Media-pouca
98	Media	Pouca	Pouca	Hll	Muita	Media-muita
99	Media	Pouca	Muita	Cir	Pouca	Pouca
100	Media	Pouca	Muita	Cir	Muita	Pouca
101	Media	Pouca	Muita	Cci	Pouca	Pouca
102	Media	Pouca	Muita	Cci	Muita	Pouca
103	Media	Pouca	Muita	Ali	Pouca	Pouca
104	Media	Pouca	Muita	Ali	Muita	Pouca
105	Media	Pouca	Muita	Ccc	Pouca	Pouca
106	Media	Pouca	Muita	Ccc	Muita	Pouca
107	Media	Pouca	Muita	Chb	Pouca	Pouca
108	Media	Pouca	Muita	Chb	Muita	Pouca
109	Media	Pouca	Muita	Chi	Pouca	Pouca
110	Media	Pouca	Muita	Chi	Muita	Pouca
111	Media	Pouca	Muita	Hll	Pouca	Media-muita
112	Media	Pouca	Muita	Hll	Muita	Media-muita
113	Média	Média	Pouca	Cir	Pouca	Pouca
114	Media	Media	Pouca	Cir	Muita	Pouca
115	Media	Media	Pouca	Cci	Pouca	Pouca
116	Media	Media	Pouca	Cci	Muita	Pouca
117	Media	Media	Pouca	Ali	Pouca	Pouca
118	Media	Media	Pouca	Ali	Muita	Pouca
119	Media	Media	Pouca	Ccc	Pouca	Pouca
120	Media	Media	Pouca	Ccc	Muita	Pouca
121	Media	Media	Pouca	Chb	Pouca	Pouca
122	Media	Media	Pouca	Chb	Muita	Pouca
123	Media	Media	Pouca	Chi	Pouca	Pouca
124	Media	Media	Pouca	Chi	Muita	Pouca
124	Media	Media	Pouca	Hll	Pouca	Media
126	Media	Media	Pouca	Hll	Muita	Media-muita
127	Media	Media	Muita	Cir	Pouca	Pouca
128	Media	Media	Muita	Cir	Muita	Pouca
129	Media	Media	Muita	Cci	Pouca	Pouca
130	Media	Media	Muita	Cci	Muita	Pouca
131	Media	Media	Muita	Ali	Pouca	Pouca
132	Media	Media	Muita	Ali	Muita	Pouca
133	Media	Media	Muita	Ccc	Pouca	Pouca

REGRA	CINZA	TAMANHO	ORIENTAÇÃO	FORMA	DISTÂNCIA	SALIÊNCIA TOPDOWN
134	Media	Media	Muita	Ccc	Muita	Pouca
135	Media	Media	Muita	Chb	Pouca	Pouca
136	Media	Media	Muita	Chb	Muita	Pouca
137	Media	Media	Muita	Chi	Pouca	Pouca
138	Media	Media	Muita	Chi	Muita	Pouca
139	Media	Media	Muita	Hll	Pouca	Media-muita
140	Media	Media	Muita	Hll	Muita	Media
141	Media	Muita	Pouca	Cir	Pouca	Pouca
142	Media	Muita	Pouca	Cir	Muita	Pouca
143	Media	Muita	Pouca	Cci	Pouca	Pouca
144	Media	Muita	Pouca	Cci	Muita	Pouca
145	Media	Muita	Pouca	Ali	Pouca	Pouca
146	Media	Muita	Pouca	Ali	Muita	Pouca
147	Media	Muita	Pouca	Ccc	Pouca	Pouca
148	Media	Muita	Pouca	Ccc	Muita	Pouca
149	Media	Muita	Pouca	Chb	Pouca	Pouca
150	Media	Muita	Pouca	Chb	Muita	Pouca
151	Media	Muita	Pouca	Chi	Pouca	Pouca
152	Media	Muita	Pouca	Chi	Muita	Pouca
153	Media	Muita	Pouca	Hll	Pouca	Media-muita
154	Media	Muita	Pouca	Hll	Muita	Media-muita
155	Media	Muita	Muita	Cir	Pouca	Pouca
156	Media	Muita	Muita	Cir	Muita	Pouca
157	Media	Muita	Muita	Cci	Pouca	Pouca
158	Media	Muita	Muita	Cci	Muita	Pouca
159	Media	Muita	Muita	Ali	Pouca	Pouca
160	Media	Muita	Muita	Ali	Muita	Pouca
161	Media	Muita	Muita	Ccc	Pouca	Pouca
162	Media	Muita	Muita	Ccc	Muita	Pouca
163	Media	Muita	Muita	Chb	Pouca	Pouca
164	Media	Muita	Muita	Chb	Muita	Pouca
165	Media	Muita	Muita	Chi	Pouca	Pouca
166	Media	Muita	Muita	Chi	Muita	Pouca
167	Media	Muita	Muita	Hll	Pouca	Media-muita
168	Media	Muita	Muita	Hll	Muita	Media-muita
169	Muita	Pouca	Pouca	Cir	Pouca	Pouca
170	Muita	Pouca	Pouca	Cir	Muita	Pouca
171	Muita	Pouca	Pouca	Cci	Pouca	Pouca
172	Muita	Pouca	Pouca	Cci	Muita	Pouca
173	Muita	Pouca	Pouca	Ali	Pouca	Pouca
174	Muita	Pouca	Pouca	Ali	Muita	Pouca
175	Muita	Pouca	Pouca	Ccc	Pouca	Pouca
176	Muita	Pouca	Pouca	Ccc	Muita	Pouca
177	Muita	Pouca	Pouca	Chb	Pouca	Pouca
178	Muita	Pouca	Pouca	Chb	Muita	Pouca
179	Muita	Pouca	Pouca	Chi	Pouca	Pouca
180	Muita	Pouca	Pouca	Chi	Muita	Pouca
181	Muita	Pouca	Pouca	Hll	Pouca	Media
182	Muita	Pouca	Pouca	Hll	Muita	Media-muita
183	Muita	Pouca	Muita	Cir	Pouca	Pouca
184	Muita	Pouca	Muita	Cir	Muita	Pouca
185	Muita	Pouca	Muita	Cci	Pouca	Pouca
186	Muita	Pouca	Muita	Cci	Muita	Pouca
187	Muita	Pouca	Muita	Ali	Pouca	Pouca
188	Muita	Pouca	Muita	Ali	Muita	Pouca

REGRA	CINZA	TAMANHO	ORIENTAÇÃO	FORMA	DISTÂNCIA	SALIÊNCIA TOPDOWN
189	Muita	Pouca	Muita	Ccc	Pouca	Pouca
190	Muita	Pouca	Muita	Ccc	Muita	Pouca
191	Muita	Pouca	Muita	Chb	Pouca	Pouca
192	Muita	Pouca	Muita	Chb	Muita	Pouca
193	Muita	Pouca	Muita	Chi	Pouca	Pouca
194	Muita	Pouca	Muita	Chi	Muita	Pouca
195	Muita	Pouca	Muita	Hll	Pouca	Media-muita
196	Muita	Pouca	Muita	Hll	Muita	Media
197	Muita	Média	Pouca	Cir	Pouca	Pouca
198	Muita	Media	Pouca	Cir	Muita	Pouca
199	Muita	Media	Pouca	Cci	Pouca	Pouca
200	Muita	Media	Pouca	Cci	Muita	Pouca
201	Muita	Media	Pouca	Ali	Pouca	Pouca
202	Muita	Media	Pouca	Ali	Muita	Pouca
203	Muita	Media	Pouca	Ccc	Pouca	Pouca
204	Muita	Media	Pouca	Ccc	Muita	Pouca
205	Muita	Media	Pouca	Chb	Pouca	Pouca
206	Muita	Media	Pouca	Chb	Muita	Pouca
207	Muita	Media	Pouca	Chi	Pouca	Pouca
208	Muita	Media	Pouca	Chi	Muita	Pouca
209	Muita	Media	Pouca	Hll	Pouca	Media-muita
210	Muita	Media	Pouca	Hll	Muita	Media-muita
211	Muita	Media	Muita	Cir	Pouca	Pouca
212	Muita	Media	Muita	Cir	Muita	Pouca
213	Muita	Media	Muita	Cci	Pouca	Pouca
214	Muita	Media	Muita	Cci	Muita	Pouca
215	Muita	Media	Muita	Ali	Pouca	Pouca
216	Muita	Media	Muita	Ali	Muita	Pouca
217	Muita	Media	Muita	Ccc	Pouca	Pouca
218	Muita	Media	Muita	Ccc	Muita	Pouca
219	Muita	Media	Muita	Chb	Pouca	Pouca
220	Muita	Media	Muita	Chb	Muita	Pouca
221	Muita	Media	Muita	Chi	Pouca	Pouca
222	Muita	Media	Muita	Chi	Muita	Pouca
223	Muita	Media	Muita	Hll	Pouca	Media-muita
224	Muita	Media	Muita	Hll	Muita	Media-pouca
225	Muita	Muita	Pouca	Cir	Pouca	Pouca
226	Muita	Muita	Pouca	Cir	Muita	Pouca
227	Muita	Muita	Pouca	Cci	Pouca	Pouca
228	Muita	Muita	Pouca	Cci	Muita	Pouca
229	Muita	Muita	Pouca	Ali	Pouca	Pouca
230	Muita	Muita	Pouca	Ali	Muita	Pouca
231	Muita	Muita	Pouca	Ccc	Pouca	Pouca
232	Muita	Muita	Pouca	Ccc	Muita	Pouca
233	Muita	Muita	Pouca	Chb	Pouca	Pouca
234	Muita	Muita	Pouca	Chb	Muita	Pouca
235	Muita	Muita	Pouca	Chi	Pouca	Pouca
236	Muita	Muita	Pouca	Chi	Muita	Pouca
237	Muita	Muita	Pouca	Hll	Pouca	Media-muita
238	Muita	Muita	Pouca	Hll	Muita	Media
239	Muita	Muita	Muita	Cir	Pouca	Pouca
240	Muita	Muita	Muita	Cir	Muita	Pouca
241	Muita	Muita	Muita	Cci	Pouca	Pouca
242	Muita	Muita	Muita	Cci	Muita	Pouca
243	Muita	Muita	Muita	Ali	Pouca	Pouca

REGRA	CINZA	TAMANHO	ORIENTAÇÃO	FORMA	DISTÂNCIA	SALIÊNCIA TOPDOWN
244	Muita	Muita	Muita	Ali	Muita	Pouca
245	Muita	Muita	Muita	Ccc	Pouca	Pouca
246	Muita	Muita	Muita	Ccc	Muita	Pouca
247	Muita	Muita	Muita	Chb	Pouca	Pouca
248	Muita	Muita	Muita	Chb	Muita	Pouca
249	Muita	Muita	Muita	Chi	Pouca	Pouca
250	Muita	Muita	Muita	Chi	Muita	Pouca
251	Muita	Muita	Muita	Hll	Pouca	Media
252	Muita	Muita	Muita	Hll	Muita	Media-pouca