Campus de São Carlos

**APLICAÇÃO DA TRANSFORMADA DE HOUGH PARA
LOCALIZAÇÃO DOS OLHOS EM FACES HUMANAS**

Autora: Lílian Saldanha Marroni

Orientador: Prof. Dr. Adilson Gonzaga

UNIVERSIDADE DE SÃO PAULO

**ESCOLA DE ENGENHARIA
DE SÃO CARLOS**

UNIVERSIDADE DE SÃO PAULO
ESCOLA DE ENGENHARIA DE SÃO CARLOS
DEPARTAMENTO DE ENGENHARIA ELÉTRICA

**APLICAÇÃO DA TRANSFORMADA DE HOUGH PARA
LOCALIZAÇÃO DOS OLHOS EM FACES HUMANAS**

LILIAN SALDANHA MARRONI

Serviço de Pós-Graduação EESC/USP

EXEMPLAR REVISADO

Data de entrada no Serviço: 12.10.91.02

Ass.:

Dissertação apresentada à escola de Engenharia de São Carlos, da Universidade de São Paulo, como parte dos requisitos para obtenção do título de Mestre em Engenharia Elétrica.

ORIENTADOR: Prof. Dr. Adilson Gonzaga

DEDALUS - Acervo - EESC

31100037089

São Carlos
2002

Class. TESE - EESC ✓
Cult. 5878
Tombo T0230/02
Sysno 1267955

Ficha catalográfica preparada pela Seção de Tratamento
da Informação do Serviço de Biblioteca – EESC/USP

M361a Marroni, Lilian Saldanha
Aplicação da transformada de Hough para localização
dos olhos em faces humanas / Lilian Saldanha Marroni.
-- São Carlos, 2002.

Dissertação (Mestrado) -- Escola de Engenharia de
São Carlos-Universidade de São Paulo, 2002.
Área: Engenharia Elétrica.
Orientador: Prof. Dr. Adilson Gonzaga.

1. Análise de imagens. 2. Processamento de imagens.
4. Transformada de Hough. 5. Distância Euclidiana.
I. Título.

FOLHA DE JULGAMENTO

Candidata: Tecnóloga **LILIAN SALDANHA MARRONI**

Dissertação defendida e julgada em 27-08-2002 perante a Comissão Julgadora:

Prof. Dr. **ADILSON GONZAGA (Orientador)**
(Escola de Engenharia de São Carlos/USP)

APROVADA

Profa. Dra. **MARIA STELA VELUDO DE PAIVA**
(Escola de Engenharia de São Carlos/USP)

APROVADA

Profa. Assoc. **AGMA JUCI MACHADO TRAINA**
(Instituto de Ciências Matemáticas e de Computação/USP)

Aprovada

Prof. Assoc. **MURILO ARAUJO ROMERO**
Coordenador do Programa de Pós-Graduação
em Engenharia Elétrica

Profa. Assoc. **MARIA DO CARMO CALIJURI**
Presidente da Comissão de Pós-Graduação

“Sê forte e corajoso; não temas, nem te espantes, pois
Eu estarei contigo por onde quer que andares.”
(Js 1:9)

A minha mãe e a minha filha, Lorena, por
estarem sempre ao meu lado.

AGRADECIMENTOS

Ao professor Adilson Gonzaga, pela orientação fornecida durante a elaboração deste trabalho.

Ao Alessandro, César, José Carlos, Rui, José Luis, Marisa e demais funcionários do Departamento de Engenharia Elétrica da EESC/SC pela disponibilidade, atenção e, principalmente, a amizade.

Aos meus amigos de São Carlos: Almir, Luciano, Rogério, Sílvio, Ana Claudia, Anali, Vivian, Viviane, César, Celso, Rodrigo, Cláudio, Maurício, Marcelo Cad, Renato e Vanessa.

Aos meus amigos de Assis: Marcos, Aline Silva, Aline Marroni, Terezinha, Ana Paula, Marcelo e Alexandre.

A Capes, pelo apoio financeiro durante o mestrado.

SUMÁRIO

LISTA DE FIGURAS.....	III
LISTA DE TABELAS	VI
LISTA DE SIGLAS.....	VII
RESUMO	VIII
ABSTRACT.....	IX
1.INTRODUÇÃO	1
1.1 ESTRUTURA DO TRABALHO	4
2. DETECÇÃO E RECONHECIMENTO DE FACES HUMANAS..	5
2.1 DETECÇÃO DE FACES	5
2.2 RECONHECIMENTO DE FACES	6
2.3 RECONHECIMENTO DE GESTOS E EXPRESSÕES	8
2.4 APLICAÇÕES	12
2.4.1 Casamento estático e dinâmico	13
3.DETECÇÃO DE OLHOS EM FACES HUMANAS	16
3.1 INTRODUÇÃO	16
3.2 TÉCNICAS PARA DETECÇÃO DE OLHOS	18
3.2.1 Segmentação por cor.....	18
3.2.2 Análise de Componentes Principais.....	21
3.2.3 Template matching	24
3.2.4 Templates deformáveis	25
3.2.5 Técnicas baseadas em Gabor	26
3.2.6 Método do Triângulo Isósceles.....	30
4. A TRANSFORMADA DE HOUGH.....	33
4.1 INTRODUÇÃO	33
4.2 FUNDAMENTOS DA TRANSFORMADA DE HOUGH	34
4.3 PRINCIPAIS MODELOS DE TRANSFORMADA DE HOUGH	50
4.3.1 Transformada de Hough Generalizada.....	50
4.3.2 Transformada de Hough Rápida	52
4.3.3. Transformada de Hough Adaptativa.....	55
5.LOCALIZAÇÃO DE OLHOS ATRAVÉS DA TRANSFORMADA DE HOUGH	58
5.1 METODOLOGIA	60
5.1.1 Pré-processamento.....	60
5.1.2 Segmentação	61
5.2 TRANSFORMADA DE HOUGH PARA OBJETOS CIRCULARES	64
5.2.1 Algoritmo da Transformada de Hough para objetos circulares	67
5.2 DISTÂNCIA EUCLIDIANA	68

6. RESULTADOS, TRABALHOS FUTUROS E CONCLUSÕES...	70
6.1 RESULTADOS	70
6.3 TRABALHOS FUTUROS E CONCLUSÕES.....	75
REFERÊNCIAS BIBLIOGRÁFICAS	77
APÊNDICES	89

LISTA DE FIGURAS

FIGURA 1.1- Pontos para extração das características de perfil	03
FIGURA 3.1- Amostra de cor da pele.....	19
FIGURA 3.2 - (a) Imagem de entrada. (b) Imagem obtida a partir do modelo de cor da pele.....	20
FIGURA 3.3 - (a) Imagem resultante da limiarização. (b) Aplicação do operador de fechamento e filtro da mediana.....	21
FIGURA 3.4 – Esquema geral para reconhecimento de faces usando PCA (GROß, 1994)	23
FIGURA 3.5 - (a) Imagem de uma face. (b) Representação obtida pela GWN. (c) Posição das 16 wavelets de maior peso.....	29
FIGURA 3.6- (a) Imagem de uma face. (b) Representações obtidas utilizando GWNs com 32, 52, 100 e 320 <i>wavelets</i> (FERIS, 2000).....	29
FIGURA 3.7 - (a) Triângulo isósceles $i j k$. (b) 3 pontos (i, j e k) formam um triângulo isósceles.....	31
FIGURA 4.1 – A Transformada de Hough. a) Espaço de imagem b) Espaço de Parâmetros.....	34
FIGURA 4.2 – (a) Representação da equação da reta em função de suas coordenadas polares. (b) Relação entre definição de reta em coordenadas polares.....	35
FIGURA 4.3– Ilustração da HT. (a) Pontos no espaço da imagem. (b) Mapeamento dos pontos no espaço $(\rho-\theta)$. (c) Detecção das propriedades de colinearidade. Ponto A denota intersecção de curvas correspondentes aos pontos 1, 3 e 5 no plano $x - y$ e similarmente ponto B corresponde aos pontos colineares 2, 3 e 4. (d) Propriedade reflexão da HT	36
FIGURA 4.4 – Forma circular contida numa imagem.....	38
FIGURA 4.5 – Os círculos paramétricos interceptam-se no centro da imagem.....	39
FIGURA 4.6 – Cruzamento entre cones no espaço de parâmetros que determinam a existência de uma forma circular na imagem.....	39

FIGURA 4.7 – Grade acumuladora quadrangular onde um grande número de círculos passa pela célula correspondente ao centro da forma circular da imagem.....	40
FIGURA 4.8– Planos constituídos de grades acumuladoras quadrangulares onde um grande número de círculos passa na célula correspondente ao centro da forma circular na imagem.....	41
FIGURA 4.9 – O centro é localizado através de um espaço acumulador bidimensional.....	46
FIGURA 4.10– O valor do raio é encontrado através de um histograma obtido a partir da medida entre os pontos de borda da imagem e o valor do centro.....	46
FIGURA 4.11- a) Imagem contendo formas circulares concêntricas; b) Espaço acumulador dos parâmetros a e b ; c) Espaço acumulador dos parâmetros a e r	48
FIGURA 4.12 - Gráfico de desempenho de cada método em relação ao número de formas circulares detectadas.....	49
FIGURA 4.13 – Usando convolução para compensação de erros na GHT.....	50
FIGURA 4.14 – Geometria da Transformada de Hough Generalizada.....	51
FIGURA 4.15 –Exemplo de um processo de FHT.....	52
FIGURA 4.16 – Detecção de segmentos lineares usando a AHT. (a) Imagem. (b) Espaço de parâmetros. (limites $m = -1$ a 1 , $c = -256$ a 256). (c) Acumulador grosseiramente quantizado.....	53
FIGURA 4.17 – Detecção do centro do círculo usando a AHT. (a) Imagem. (b) Espaço de parâmetros (limites $a = 0$ a 128 , $b = 0$ a 128). (c) Acumulador grosseiramente quantizado.....	56
FIGURA 4.18 – Detecção do centro do círculo usando a AHT. (a) Imagem. (b) Espaço de parâmetros (limites $a = 0$ a 128 , $b = 0$ a 128). (c) Acumulador grosseiramente quantizado.....	57
FIGURA 5.1- Fluxograma do software desenvolvido.....	59
FIGURA 5.2 – Imagem adquirida da Internet.....	60
FIGURA 5.3 – Exemplo de filtro passa-baixa no domínio do espaço.....	61
FIGURA 5.4 - Máscaras utilizadas no detector de Sobel.....	61
FIGURA 5.5 - Determinação da Magnitude dos gradientes e também do ângulo.....	62

FIGURA 5.6 – (a) Imagem de entrada; (b) resultado do filtro da média 3x3; (c) resultado de Sobel	63
FIGURA 5.7 – Imagem com uma circunferência	65
FIGURA 5.8 – Plano de parâmetros “a” e “b” para circunferências	66
FIGURA 5.9 – Circunferência geradas com centro em <i>pixels</i> da imagem que ficam nas bordas.....	66
FIGURA 5.10 - Dimensões do arranjo acumulador.....	66
FIGURA 5.11 - (a) Imagem de entrada; (b) Arranjo acumulador criado pela HT	68
FIGURA 5.12 – Proporção do cálculo da Distância Euclidiana entre os olhos e a largura da imagem.....	69
FIGURA 5.13- Localização dos olhos pela proporção de 33%.....	69
FIGURA 6.1 - Resultado da detecção de olhos com variação de expressão.....	70
FIGURA 6.2- Resultado da detecção de olhos com variação de rotação.....	71
FIGURA 6.3- Resultado da detecção de olhos com olhos entreabertos e fechados.....	71
FIGURA 6.4 : (a) Imagem de entrada; (b) conversão da imagem para binária; (c) operação de abertura; (c) operação de fechamento.....	72
FIGURA 6.5 - Os pontos vermelhos representam os centros de cada bloco (ampliados 8x).....	72
FIGURA 6.6 - Detecção dos olhos através da geometria do triângulo isósceles.....	73
FIGURA 6.7 - Imagens com problemas na localização dos centros dos blocos	73
FIGURA 6.8: (a) Imagem binarizada; (b) erosão e (c) dilatação para realização da operação de abertura.....	74

LISTA DE TABELAS

TABELA 1.1 - Características para identificar faces de perfil (SAMAL & IYENGAR, 1992)	02
TABELA 1.2 - Características para identificar faces frontais (SAMAL & IYENGAR, 1992)	03
TABELA 2.1 – Unidades de Ação do Sistema de Código de Ação Facial (ZLOCHOWER, 1997).....	09
TABELA 2.2 – Ações variadas (ZLOCHOWER, 1997).....	10
TABELA 2.3 - Aplicações da Tecnologia de Reconhecimento de Faces (CHELLAPPA, 1996).....	12
TABELA 4.1 – Efeito da HT sobre uma linha reta no espaço m-c.....	34
TABELA 4.2 – Valores referentes ao exemplo do processo de FHT	54
TABELA 6.1- Testes comparativos entre os métodos utilizados.....	74

LISTA DE SIGLAS

2-IHT	Transformada de Hough 2-1
AHT	Transformada de Hough Adaptativa
AU	Action Units
FACS	Facial Action Coding System
FHT	Transformada de Hough Rápida
FRT	Face Recogniton Tecnology
GHT	Transformada de Hough de Gerig
GHTG	Transformada de Hough de Gerig com Gradiente
GWN	Gabor Wavelet Network
GWT	Gabor Wavelet Template
HT	Transformada de Hough
ISFER	Integrated System for Facial Expression Recognition
KL	Karhunen-Loève
MLP	Multi-Layer Perceptron
PCA	Principal Component Analysis
RBF	Radial Basis Function
SHT	Transformada de Hough Padrão

RESUMO

Com a crescente necessidade de segurança, o processo de identificação pessoal é cada vez mais exigido. A extração de características faciais é um passo importante quando se lida com interpretação visual automatizada no reconhecimento de faces humanas. Dentre as características faciais, os olhos são partes importantes no processo de reconhecimento, pois determinam o início da busca por outras características relevantes. Neste trabalho é apresentado um método de localização de olhos em imagens frontais de faces humanas. Este método é subdividido em duas partes. Primeiro, são identificados os possíveis candidatos a olhos usando a Transformada de Hough para círculos; depois é aplicada a Distância Euclidiana confirmando-se a localização por biometria facial.

Palavras-chave: Análise de Imagens; Processamento de Imagens; Transformada de Hough; Distância Euclidiana.

ABSTRACT

Personal identification process is an exigency for security systems. Facial feature extraction is a crucial step for automated visual interpretation in human face recognition. Within all the facial features, the eyes are significantly parts for the recognition process, therefore they set up the start for another relevant feature search. In this work, we present a method for eyes locating in digital images of frontal human faces. This method is subdivided into two parts. First, we identify the possible eyes's candidates by Hough Transform for circles; then we apply the Euclidian Distance and calculate the eye's position by facial biometric measurement.

Key-words: Image analysis; Image processing; Hough Transform; Euclidean Distance.

Capítulo 1

Introdução

Sistemas para localizar e identificar faces humanas são úteis em várias aplicações: são utilizados em teleconferências, vídeo-fone e sistemas de segurança, entre outros. Embora seja claro que as pessoas consigam reconhecer facilmente as faces, não é tão claro como as mesmas são codificadas e armazenadas e muito menos como a familiaridade afeta este processo complexo (HUANG *et al.*, 1998). Na prática, verifica-se que não é fácil implantar a experiência ou conhecimento para reconhecimento de faces humanas em máquinas. Nos últimos vinte anos, extensivas pesquisas têm sido conduzidas por psicofísicos, neuro-cientistas e engenheiros em vários aspectos do reconhecimento de faces por humanos e máquinas. Os dois primeiros têm se preocupado com questões como: unicidade de faces; se o reconhecimento é feito holisticamente ou pela análise de característica local; análise e uso de expressões faciais para reconhecimento; como crianças percebem as faces; organização de memória por faces; incapacidade para reconhecer precisamente faces invertidas, existência de um neurônio “*grandmother*” para reconhecimento de faces, incapacidade de reconhecimento devido a condições como prosopagnosia; o papel do hemisfério direito do cérebro em percepção de faces. Muitas das teorias e hipóteses desenvolvidas adiante por pesquisadores nestas disciplinas têm sido baseadas em um conjunto de faces bastante pequeno. No entanto, várias das descobertas têm conseqüências importantes para cientistas que desenvolvem algoritmos e sistemas para reconhecimento de faces humanas.

Durante o início e meados de 1970, técnicas típicas de classificação de padrões, que usam medidas de atributos entre características de face ou perfis de faces, foram usadas. Desde o início de 1990, pesquisas em FRT (*Face Recognition Technology*) apresentaram um crescimento significativo. Atribui-se a isso algumas razões: aumento nas pesquisas de projetos comerciais e civis, o ressurgimento de classificadores e redes neurais com ênfase em adaptação e computação em tempo real, aumento da necessidade de aplicações em segurança devido ao tráfico e terrorismo.

Nos últimos cinco anos, cresceu o cuidado com problemas como segmentação e localização de uma face numa dada imagem e a extração de características como olhos, boca, nariz etc. Também vários avanços foram obtidos no desenvolvimento de classificadores estatísticos e de redes neurais para faces, além dos conceitos clássicos, como métodos baseados na Transformada de KL e *matching*.

A extração de características faciais desempenha um papel importante na interpretação visual e no reconhecimento de faces, sendo características faciais extraídas do perfil ou frontais. Do perfil são facilmente extraídas, mas deve haver uma boa separação entre o fundo e a face, e esta última não deve sofrer rotação para que se obtenha uma extração confiável, como mostra a Figura 1.1 (LEE *et al.*, 1996).

Basear um procedimento de extração de características e posterior reconhecimento somente em visões frontais ao invés de visões laterais é uma boa escolha, já que uma visão frontal pode conter um conjunto mais rico de características, além da maioria dos documentos registrarem imagens de faces somente com visões frontais, tais como passaporte e registro geral (RG). SAMAL & IYENGER (1992) fornecem várias características que podem ser extraídas tanto do perfil (Tabela 1.1) quanto da frente da face (Tabela 1.2).

TABELA 1.1 - Características para identificar faces de perfil (SAMAL & IYENGAR, 1992).

CATEGORIA	CARACTERÍSTICAS
Geral	altura, contorno, “ <i>template</i> ”, área, centróide
Nariz	Protuberância do nariz (distância da linha base 2-3 ao nariz 1)
Queixo	Protuberância do queixo (distância da linha 6-10 ao queixo 2)
Ângulo	1-2-3, 1-3-4, 4-1-5
Área	Direita da linha base 2-3 Direita da linha base 6-10
Razão*	$D(1,4) / D(1,2)$
Distâncias*	$D(1,2)$, $D(1,4)$, $D(1,5)$, $D(1,8)$, $D(3,4)$, $D(4,2)$, $D(5,10)$, $D(7,9)$
Outras*	Coefficiente de correlação, momentos invariantes

* $D(x,y)$ é a distância entre as características x e y.

FIGURA 1.1 - Pontos para extração das características de perfil.

TABELA 1.2 - Características para identificar faces frontais (SAMAL & IYENGAR, 1992).

CATEGORIA	CARACTERÍSTICAS
Geral	Altura, contorno, <i>template</i> , área, centróide
Olhos	<i>template</i> , forma, cor/granulação, áreas, abertura, centro, intensidade ao redor da pupila, D (pálpebra superior, centro do olho), distâncias entre cantos, distâncias entre centro dos olhos
Pálpebras	área, comprimento
Sobrancelhas	espessura, D (olhos, sobrancelha)*
Lábios	espessura, granulação, <i>template</i> , largura
Boca	<i>template</i> , largura, altura, área, área de abertura
Nariz	<i>Template</i> , largura, comprimento, granulação, área da narina
Orelha	comprimento, granulação, área
Cabelo	Intensidade/granulação/cor, quantidade/cobertura, comprimento, textura
Bochechas	Intensidade
Distâncias*	largura da cabeça na altura dos olhos, D (queixo, linha dos olhos), D (centro dos lábios, queixo), D (olhos, linha do cabelo), D (olhos, centro do nariz), D (extremidade do queixo, centro da face), D (borda esquerda, centro do nariz), D (boca direita, centro de nariz), D (canto interno do olho, centro da face), D (centro do olho, centro da testa), D (centro da face, centro da sobrancelha), D (canto interno do olho, centro da boca), D (centro da boca, centro do olho), D (centro da boca, fim do nariz)
Áreas	cantos internos dos olhos e centro da boca centro da pálpebra, centro do olho e centro da face centro da boca e centro das sobrancelhas
Razão*	D (centro da face, abertura da boca) / D (queixo, abertura da boca) D (queixo, abertura da boca) / D (fim do nariz, abertura da boca)
Outros	contorno do queixo, tamanho da testa, contorno da testa

*D (x,y) é a distância entre as características x e y.

1.1 Estrutura do trabalho

No capítulo 2 são discutidos alguns aspectos do processamento de faces, tais como a detecção, que é um pré-processamento necessário para qualquer sistema de reconhecimento de faces, o reconhecimento, propriamente dito, suas aplicações e o reconhecimento de gestos e expressões faciais.

No terceiro capítulo é apresentada a importância da detecção de olhos e uma das etapas fundamentais em tarefas de processamento digital de imagens: a segmentação, bem como alguns extratores de características. A discussão é um pouco extensa, principalmente porque o último algoritmo foi implementado para realização de testes comparativos com o algoritmo proposto nesta dissertação.

No quarto capítulo são mostradas as técnicas descritas na literatura para implementação da Transformada de Hough.

No quinto capítulo discute-se a abordagem proposta para a detecção de olhos em imagens de faces humanas.

No último capítulo são apresentados os resultados obtidos e indicados trabalhos futuros para um maior desenvolvimento do sistema.

Capítulo 2

Detecção e Reconhecimento de Faces Humanas

2.1 Detecção de Faces

A Detecção Automática de Faces humanas constitui um problema complexo no desenvolvimento e aplicações de Reconhecimento de Padrões (YANG & HUANG, 1994), o qual merece especial estudo, uma vez que é o pré-processamento necessário para qualquer sistema de Reconhecimento de Faces.

Para possibilitar um reconhecimento preciso, primeiramente, é necessário detectar as faces e posteriormente segmentá-las para que somente essas sejam tratadas, não considerando os objetos que estiverem próximas do sujeito a ser reconhecido, os quais podem influenciar a tomada de decisão nas demais etapas do processo.

A segmentação automática deve localizar a região da face, com seus respectivos componentes: olhos, boca, nariz etc, distinguindo a face humana a partir de outros tipos de padrões no fundo.

Três fatores tornam esse problema mais complexo. Primeiro, variações de padrões devido a diferenças na aparência facial, discrepâncias entre imagens diferentes de uma mesma face, expressão e cor da pele. Segundo, ocultação de características faciais por objetos como óculos e bigode e também cabelo na face e maquiagem. Terceiro, a escala e a orientação da face na imagem, o que dificulta a utilização de *templates* fixos para localizar as características procuradas. Outras dificuldades encontradas devido à presença de ruídos são oclusões e mudanças na iluminação, as quais podem projetar ou remover sombras (SUNG & POGGIO, 1998).

Duas são as abordagens principais para a Detecção de Faces. Na primeira, a face é tratada atômica e um modelo geral representando os principais traços da face é utilizado como aproximação para a face que se deseja localizar. Esse modelo pode ser estático (quando informações prévias são conhecidas de antemão, por exemplo: o tamanho aproximado das faces em relação à imagem total ou o número

de faces presentes na cena) ou dinâmico (métodos mais genéricos em que nenhuma informação a respeito da cena é previamente conhecida).

Na segunda abordagem, a face é localizada através de alguns de seus componentes, tais como olhos, boca e nariz. Da mesma forma que na abordagem anterior, a disponibilidade de informações prévias definirá uma estratégia mais específica ou genérica.

2.2 Reconhecimento de Faces

O Reconhecimento de Faces é uma área muito promissora, que ainda possui diversos problemas não resolvidos, permanecendo como uma das mais desafiadoras áreas no campo da Visão Computacional e do Reconhecimento de Padrões.

Um considerável esforço tem sido dedicado à resolução do problema de reconhecimento e caracterização de objetos e faces presentes em uma imagem. No início dos anos 80, com os avanços alcançados pela microeletrônica e no desenvolvimento de arquiteturas paralelas de processamento, as técnicas digitais passaram a ser utilizadas para esse fim (PERELMUTER & CARRERA, 1995).

Devido a sua importância prática e aos interesses dos cientistas cognitivos, a pesquisa em Reconhecimento de Faces é tão antiga quanto a própria Visão Computacional (PENTLAND, 2000). CHELLAPPA *et al.* (1995) apresentam uma análise de trinta anos de pesquisa em reconhecimento de faces humanas, citando 221 trabalhos. Atualmente, a melhoria das técnicas computacionais inteligentes como Sistemas Especialistas, Lógica Nebulosa, Redes Neurais, Algoritmos Genéticos e PCA, têm permitido novas abordagens para esse problema, com mais precisão e maior abrangência de aplicação.

Métodos seguros e robustos podem ser usados em uma grande variedade de aplicações, como por exemplo, em sistemas de segurança de empresas e aeroportos ou em sistemas de identificação de suspeitos (GROß, 1994).

Outro sistema que vem sendo desenvolvido é o de identificação de pessoas mesmo que essas não percebam que estão sendo observadas, de forma que no futuro, possivelmente, uma criança desaparecida poderá ser localizada através de imagem de

câmeras espalhadas em pontos estratégicos de uma cidade, como estações de metrô, cruzamento de avenidas ou até mesmo obtidas por meio de satélites.

Além dessas aplicações, várias outras que motivam a pesquisa nessa área são analisadas por (CHELLAPPA *et al.*, 1995), como:

- identificação pessoal para banco, passaporte, fichas criminais;
- sistemas de segurança e de controle de acesso;
- monitoramento de multidões em estações, shopping centers, etc;
- criação de retrato falado;
- busca em fichas criminais;
- envelhecimento computadorizado para auxiliar na busca de desaparecidos;
- interfaces perceptuais homem-máquina com Reconhecimento de Expressões Faciais.

Os softwares para reconhecimento de faces impõem algumas condições para funcionamento correto. LEE *et al.* (1996) assumem as seguintes condições para extrair as características:

- a face deve estar localizada no centro da imagem;
- os olhos devem estar abertos e a boca fechada;
- o ângulo de rotação da imagem deve ser menor que 5°;
- a face não deve estar escondida por sombras, óculos, bigode ou barba.

Um sistema de Reconhecimento de Faces deve abordar diversas questões técnicas relativas tanto ao estímulo tratado (a face humana) quanto aos processos envolvidos na tarefa de reconhecimento em si.

São vários os processos associados ao Reconhecimento Automático de Faces (PERELMUTER, 1995):

- **Detecção:** dada uma imagem qualquer, identifica o local onde aparece(m) a(s) face(s);
- **Segmentação:** identifica as partes constituintes de uma face;
- **Representação:** seleciona as informações da face que serão utilizadas para representá-la;
- **Identificação:** verifica se uma face pertence ao conjunto de faces “conhecidas”.

2.3 Reconhecimento de gestos e expressões

Recentemente, a análise de gestos e expressões faciais tem atraído atenção em Visão Computacional. Um gesto pode ser definido como sendo o movimento do corpo, especialmente da cabeça e dos braços, com a finalidade de exprimir idéias, sinais ou mímicas. Eles variam desde simples ações, como apontar objetos, até movimentos mais complexos, que expressam sentimento e permitem a comunicação com outras pessoas.

O Reconhecimento de Gestos é uma área de pesquisa muito ativa. HUANG & PAVLOVIC (1995) apresentam uma revisão de trabalhos relacionados com a modelagem, análise e síntese de gestos de mão. Alguns trabalhos anteriores foram apresentados por BAUDEL & BEAUDOUIN-LAFON (1993), que usaram uma luva mecânica para controlar a apresentação no computador. FUKOMOTO *et al.* (1992) também apresentaram um método para guiar a apresentação no computador, mas sem usar uma luva. DARELLE & PENTLAND (1993) usaram modelos de vista, os quais são automaticamente treinados através de uma seqüência de imagens para representar todas as possíveis posições de uma mão, usando correlação.

A maior parte dos sistemas automáticos de análise de expressões tenta reconhecer um pequeno grupo de expressões: alegre, surpreso, irado, triste, medroso e com nojo. Tais expressões, entretanto, não ocorrem freqüentemente. As emoções humanas e intenções são comunicadas mais vezes pelas mudanças discretas em uma ou duas características faciais, como apertar os lábios na ira ou abaixar os seus cantos na tristeza (KANADE *et al.*, 2000a).

EKMAN & FRIESEN (1978) desenvolveram um Sistema de Código de Ação Facial (FACS) para descrever mudanças sutis nas expressões faciais. Este sistema consistia de 44 unidades de ação (AU), incluindo posições do olho e da cabeça. Trinta dessas eram anatomicamente relacionadas para contração de um conjunto específico de músculos faciais. A Tabela 2.1 mostra as AU e a Tabela 2.2 as ações variadas. A base anatômica das outras 14 não era especificada, sendo referida como ações variadas.

Embora exista um pequeno número de unidades de ações anatômicas, mais de 7.000 combinações de AU têm sido observadas.

TABELA 2.1– Unidades de Ação do Sistema de Código de Ação Facial (ZLOCHOWER, 1997).

AU	Músculo facial	Descrição do movimento do músculo
1	Frontal, parte medial	Canto interno da sobrancelha elevado
2	Frontal, parte lateral	Canto externo da sobrancelha elevado
4	Corrugador do supercílio, depressor do supercílio	Sobrancelhas desenhadas medialmente e para baixo
5	Levantador da pálpebra superior	Olhos ampliados
6	Orbicular do olho, parte orbital	Bochechas elevadas, olhos estreitos
7	Orbicular do olho, parte palpebral	Pálpebra inferior elevada e desenhada medialmente
9	Levantador do lábio superior e da asa do nariz	Lábio superior elevado e invertido, parte superior do sulco nasolabial aumentado
10	Levantador do lábio superior	Lábio superior elevado, sulco nasolabial aumentado
11	Levantador do ângulo oral (a.k.a. canino)	Parte inferior a medial do sulco nasolabial aumentada
12	Zigomático maior	Cantos dos lábios puxados para cima e lateralmente
13	Zigomático menor	Ângulo da boca elevado
14	Bucinador	Cantos dos lábios pressionados. Bochechas comprimidas contra os dentes
15	Depressor do ângulo oral (a.k.a. triangular)	Canto da boca puxado para baixo e para dentro
16	Depressor do lábio inferior	Lábio superior puxado para baixo e lateralmente
17	Mento	Pele do queixo elevada
18	Incisivo do lábio superior e incisivo do lábio inferior	Lábios franzidos
20	Risório/platisma	Cantos dos lábios puxados lateralmente
22	Orbicular da boca	Lábios afinilados
23	Orbicular da boca	Lábios contraídos
24	Orbicular da boca	Lábios pressionados juntos
25	Depressor do lábio inferior ou relaxador do mento ou orbicular da boca	Lábios afastados
26	Masseter, relaxador temporal e pterigóideo interno	Maxilar diminuído
27	Pterigóideo e digástrico	Boca esticada aberta
28	Orbicular da boca	Lábios succionados
41	Relaxador do levantador da pálpebra superior	Pálpebra superior caída
42	Orbicular do olho	Olhos semiabertos
43	Relaxador do levantador da pálpebra superior; orbicular do olho, parte palpebral	Olhos fechados
44	Orbicular do olho, parte da pálpebra	Olhos semiabertos
45	Relaxador do levantador da pálpebra superior, orbicular do olho, parte palpebral	Piscar
46	Relaxador do levantador da pálpebra superior, orbicular do olho, parte palpebral	Piscar

TABELA 2.2 – Ações variadas (ZLOCHOWER, 1997).

AU	Descrição do Movimento
8	Aproximar os lábios
19	Mostrar a língua
21	Apertar o pescoço
29	Apertar o maxilar
30	Movimentar a mandíbula para o lado
31	Cerrar a mandíbula
32	Morder os lábios
33	Soprar
34	Tragar
35	Comprimir as bochechas
36	Sobressair a língua
37	Esfregar os lábios
38	Dilatar o nariz
39	Comprimir o nariz

Em 1997 e 1998, cada um dos quatro mais influentes grupos de pesquisadores no campo de análise de gestos faciais baseados em visão publicaram um sumário de seus trabalhos.

ESSA & PENTLAND (1997) apresentaram os resultados do reconhecimento e classificação de emoção singular de expressões faciais através de um método de fluxo ótico associado à geometria e modelos físicos de face baseados em movimento. Eles usaram energia de movimento 2D e *templates* que codificaram ambos - a magnitude e a direção do movimento. Consideraram quatro expressões de emoção: alegria, ira, tristeza e surpresa, definindo *templates* no espaço temporal para todas. Embora a abordagem proposta por Essa e Pentland não tenha sido ainda validada completamente, deveria ser notado que os *templates* de espaço temporal de expressões faciais formam um método único de classificação de emoção de expressão facial.

BLACK & YACOOB (1998) também utilizaram um modelo de fluxo ótico de movimento da imagem para analisar as expressões faciais. O trabalho explorou o uso de modelos de fluxo ótico parametrizado para o reconhecimento de seis expressões de emoções básicas: alegre, surpreso, irado, triste, medroso e com nojo.

KOBAYASHI & HARA (1997) apresentaram reconhecimento em tempo-real, classificação de emoção singular e combinação de seis expressões básicas. Eles trabalharam na realização de uma face de robô 3D animada que podia reconhecer e

reproduzir as expressões. Usaram dados de distribuição de brilho de imagens faciais e rede neural *backpropagation* de 3 camadas para a classificação e síntese das expressões faciais.

Os pesquisadores do MIRALab, THALMANN *et al.* (1998), implementaram o reconhecimento, a classificação emocional singular e a animação de expressões de faces humanas utilizando um modelo virtual 3D.

Cada um dos métodos descritos anteriormente possuem algumas limitações por usarem alguns procedimentos semi-automáticos, ou completamente manuais para rastrear características faciais.

O reconhecimento automático de AU é um problema difícil porque não existem definições quantitativas e as mesmas aparecem em combinações complexas (KANADE *et al.*, 2000b). A partir de 1999, pesquisadores têm se empenhado arduamente para a criação de sistemas totalmente automáticos.

BARTLETT *et al.* (1999) construíram um sistema híbrido para reconhecer 6 AUs em vista superior da face, usando 50 projeções, 5 medidas de características e 6 templates de movimento.

KANADE *et al.* (2000b) desenvolveram um sistema de análise automática de AU usando características faciais. Inicialmente, a orientação da cabeça é detectada e a partir daí mudanças sutis nos componentes faciais são medidas. As AUs são usadas como uma coleção de parâmetros faciais e classificadas por uma rede neural. Nesse trabalho, foi focalizado o reconhecimento em AUs de faces ligeiramente inclinadas olhando para baixo porque os lábios são muito mais deformáveis que os olhos. Nove parâmetros são usados para descrever o formato, o movimento e o estado dos lábios e sulcos das faces.

SILVA & NG (2000) descreveram o uso de técnicas estatísticas e *Hidden Markov Models* (HMM) no reconhecimento de expressões faciais (vídeo) e emoções na fala (áudio).

PANTIC & ROTHKRANTZ (2000) desenvolveram um Sistema para o Reconhecimento de Expressões Faciais (ISFER). Esse sistema fornecia informações quantitativas e qualitativas sobre diferentes sinais de comunicação não-verbal em diferentes níveis. No mais baixo nível, o sistema detectava um sinal não-verbal mostrado pela pessoa observada. No próximo nível, o sistema categorizava o sinal

detectado como uma ação facial específica (por exemplo, sorrir), uma ação corporal específica (por exemplo, encolher os ombros), uma reação vocal específica (por exemplo, falar rápido), ou uma reação fisiológica específica (por exemplo, suar). No mais alto nível, o sistema deveria reconhecer as intenções do usuário.

2.4 Aplicações

A tabela 2.3 mostra algumas aplicações de reconhecimento de faces:

TABELA 2.3 - Aplicações da Tecnologia de Reconhecimento de Faces (CHELLAPPA *et al.*, 1995).

APLICAÇÕES	VANTAGENS	DESVANTAGENS
1a. Cartões de crédito, Carteira de motorista, Passaporte e Identificação pessoal	Imagem controlada Segmentação controlada Boa qualidade das imagens	Banco de dados muito grande
1b. Casamento (“ <i>matching</i> ”) de fotos de rosto	Qualidade da imagem misturada Mais de uma imagem disponível	Baixa qualidade da imagem Segmentação descontrolada
2. Segurança de bancos e lojas	Alta confiabilidade Pesquisa geograficamente localizada	Baixa qualidade da imagem Segmentação descontrolada
3. Vigilância de multidão	Alta confiabilidade Arquivos pequenos Disponibilidade de imagens de vídeo	Tempo-real Segmentação descontrolada Baixa qualidade da imagem
4. Perícia	Alta confiabilidade	Baixa qualidade da imagem
5. Retrato Falado	Pesquisa limitada	Similaridade desconhecida
6. Álbum Eletrônico de Suspeitos	Descritores de pesquisa limitados	Cansaço do examinador
7. Montagem por fotografia	Descritores de pesquisa limitados	Cansaço do examinador
8. Reconstrução de faces a partir de fragmentos fotográficos	Alta confiabilidade	Requer exemplos de entrada
9. Envelhecimento computadorizado	Alta confiabilidade	Requer exemplos de entrada

As aplicações comerciais e jurídicas da FRT (Face Recognition Technology) listadas na tabela 2.3 colocam uma grande extensão de desafios tecnicamente diferentes e requerem, igualmente, técnicas de processamento de imagem, análise,

interpretação e Reconhecimento de Padrões. Esses desafios e técnicas podem ser classificados dentro de dois grupos: casamento (*matching*) estático e casamento dinâmico (vídeo). Apesar de ambos os grupos possuírem a maioria das imagens com fundo desordenado, colocando desafios aos algoritmos de segmentação, entre eles existem diferenças significativas em termos de qualidade de imagem, disponibilidade de um critério de *matching* bem definido, natureza, tipo e a quantidade de dados de entrada de um examinador, como nas aplicações 4 e 5.

Três diferentes problemas surgem nas aplicações da Tabela 2.3: *matching*, detecção de similaridade e transformação. As aplicações 1, 2 e 3 envolvem *matching* de uma imagem facial com uma outra. As aplicações 4, 5, 6 e 7 envolvem a descoberta ou a criação de uma imagem facial que seja similar às lembranças humanas de uma face. Finalmente, as aplicações 8 e 9 envolvem a reconstrução de face a partir de uma imagem de entrada.

Cada uma dessas aplicações impõe diferentes requerimentos no processo de reconhecimento. O *matching* requer que o candidato à imagem facial esteja em algum conjunto de imagens faciais selecionado pelo sistema. A detecção de similaridade requer, em adição ao *matching*, que imagens de faces sejam encontradas. Essas imagens são similares a uma face recuperada e requerem que a medida de similaridade usada pelo sistema de reconhecimento minuciosamente combine com as medidas de similaridade usadas pelos humanos. As transformações requerem que novas imagens criadas pelo sistema sejam similares às lembranças humanas de uma face.

2.4.1 Casamento estático e dinâmico

No **casamento estático**, o casamento de fotos de rosto é a mais comum aplicação neste grupo e para simplificar a segmentação e os algoritmos de casamento, alguns cuidados no processo de aquisição das imagens devem ser tomados. Esses cuidados, já considerados padrões, envolvem o tipo de fundo, iluminação, resolução da câmera e a distância entre a câmera e a pessoa a ser fotografada. Aplicações usuais do problema de casamento de fotos de rosto são o

Reconhecimento de Faces em carteiras de motoristas, cartões de créditos, cartões de identificação pessoal e passaportes. As imagens desses documentos são normalmente adquiridas com mais controle do que em outras.

Grandes desafios surgem na procura de um banco de dados extenso e, embora as condições de imagens sejam controladas, variações na face devido ao envelhecimento, calvície, crescimento do cabelo, disfarces etc., contribuem para dificultar a extração de características e o casamento das imagens.

A aplicação 2 é mais complexa que a 1 devido ao descontrole das condições de aquisição de imagens. A baixa qualidade das imagens e a presença de um possível fundo desordenado tornam a segmentação mais difícil.

As aplicações 4, 5, 6 e 7 envolvem a descoberta ou a criação de uma imagem facial similar às lembranças humanas. Na aplicação 4 uma lista de faces similares é gerada usando um algoritmo de identificação facial. Uma perícia analisa cuidadosamente essas imagens e confirma se uma face numa dada imagem corresponde à pessoa em questão.

Na aplicação 5, é solicitado à testemunha que desenhe o acusado usando uma biblioteca de características com vários formatos de nariz, lábios, olhos, etc.

Nas aplicações 6 e 7, uma testemunha identifica uma face a partir de um conjunto de imagens faciais. Nessas aplicações, a qualidade da imagem tende a ser baixa e a similaridade difícil de quantificar. A dificuldade aumenta quando se procuram fotos em álbuns: uma tarefa desgastante que demanda longo tempo e confunde a testemunha à medida que ela examina milhares de fotos. A probabilidade de selecionar um suspeito correto (acerto) diminui rapidamente depois das primeiras 100 a 200 fotos, enquanto que a probabilidade de se selecionar uma pessoa errada (alarme falso) aumenta rapidamente (LEE, 1993).

Em resposta a tais dificuldades, têm crescido as tentativas de recuperação de imagens através de características faciais. Um conjunto de características é usado para se distinguir um subconjunto de fotos de um álbum numa base de dados. Nos sistemas de recuperação de características existentes, as testemunhas descrevem um suspeito, por exemplo, através de uma foto de alguém que se deseja encontrar ou em resposta a uma série de perguntas. Um exemplo de pergunta seria “Qual a altura dessa pessoa?” e a resposta constaria numa escala de 5 pontos, sendo baixo 1 e alto,

5. O sistema combina a descrição da testemunha com aquelas das fotos do álbum na base de dados. As fotos são apresentadas para a testemunha ordenadas por grau de similaridade.

As aplicações 8 e 9 envolvem transformações de imagens a partir de uma atual para o que elas poderiam ter sido (aplicação 8) ou para o que elas serão (aplicação 9). Tais aplicações são mais difíceis que as 4, 5 e 6.

No **casamento dinâmico**, empregado nas aplicações 2 e 3, as imagens disponíveis através de uma câmera de vídeo tendem a ser de baixa qualidade e, no caso da aplicação 3, tem-se um fundo desordenado. Uma das fortes restrições desta aplicação é a necessidade de reconhecimento em tempo real.

Capítulo 3

Detecção de Olhos em Faces Humanas

3.1 Introdução

Com a crescente necessidade de segurança, o processo de identificação pessoal é cada vez mais exigido e a detecção de características faciais é uma opção importante para a interpretação visual automatizada (HUANG *et al.*, 1998). Dentre as diversas características faciais disponíveis, os olhos representam o mais importante papel em normalização de faces e facilita, posteriormente, a localização de outros elementos faciais (BRUNELLI & POGGIO, 1993), (LAM & YAN, 1998), (HAN *et al.*, 2000), (KAWAGUCHI *et al.*, 2000), (SIROHEY & ROSENFELD, 2001), (HUANG *et al.*, 1998), (TOLBA *et al.*, 2000), (LAM & LI, 1998), (YUILLE *et al.*, 1992), (SMERALDI *et al.*, 2000), (WANG & TAN, 2000), (FUKUI & YAMAGUCHI, 1997). Como as posições dos olhos e a distância interocular são relativamente constantes para a maioria das pessoas, a detecção dos olhos provê uma excelente conjuntura para a normalização de imagens e restringe a pesquisa por outros componentes faciais. Segundo (KAWAGUCHI *et al.*, 2000), quando o tamanho, a localização e a orientação de uma face em uma imagem podem ser variados, esta face deve ser normalizada para um tamanho padrão usando as posições dos olhos. Se as características detectadas forem três (incluindo os olhos), a rotação da cabeça também pode ser normalizada.

Existem duas abordagens para a detecção automática de olhos (HUANG *et al.*, 1998):

Abordagem holística: os métodos holísticos consideram todos os *pixels* da imagem ou de regiões características da face. Nessa abordagem, a dimensionalidade dos dados é igual ao número de *pixels* das imagens consideradas. Para resolver esse problema podem ser utilizados métodos de redução de dimensionalidade, como PCA, discriminantes lineares ou redes neurais. O método de extração com PCA é o mais popular, tendo sido utilizado várias vezes em associação com pré-processamento de

normalização de imagens para melhorar o desempenho. A classificação pode ser feita de diversas maneiras, geralmente através de redes neurais ou sistemas estatísticos.

Em BRUNELLI & POGGIO (1993) é feito um estudo comparando o desempenho de métodos por atributos (locais ou abstrativos) com um método holístico. O método holístico usado foi o template matching, o qual testou o desempenho de imagens de olhos, narizes, bocas e também de toda a face. Os métodos holísticos proporcionam resultados melhores que os locais. A vantagem de abordagens globais está no fato de que pequenas variações locais pouco prejudicam a detecção, de uma forma geral. A principal desvantagem está nos problemas de variação de iluminação e, em alguns casos, no custo computacional.

Abordagem abstrativa: essa abordagem é bastante intuitiva. Alguns métodos automáticos de extração dessa categoria baseiam-se no contorno: elaboram um vetor de características a partir de medidas de distância e ângulos entre os pontos característicos da face, como pontos dos cantos dos lábios, centro da boca, nariz, narinas, pupilas, pontos extremos dos olhos, sobrancelhas, orelhas, pontos do contorno do queixo, etc. e combinação entre distâncias. YUILLE *et al.* (1992) descreve uma estratégia complexa, mas genérica, para localizar características faciais, tais como os olhos, usando *templates* deformáveis. Essa técnica será discutida com detalhes no item 3.2.4.

A detecção dos olhos também pode ser efetuada utilizando suas características simétricas. Como exemplo, YESCHURUN *et al.*, (1991) desenvolveram um operador simétrico generalizado para detecção de olhos, enquanto que DUCOTTET *et al.* (1994) detectaram os olhos com simetria circular e tamanho variável usando a Transformada Wavelet.

A vantagem dessa abordagem é a invariância à translação, e a desvantagem está nos problemas ocasionados com alterações devido a expressões faciais e a rotações.

3.2 Técnicas para Detecção de Olhos

3.2.1 Segmentação por cor

A segmentação é uma das etapas fundamentais em tarefas de processamento digital de imagens que consiste em subdividir a imagem em suas partes (ou objetos) constituintes. A segmentação automática é considerada uma das mais difíceis tarefas de processamento digital de imagens (GONZALES & WOODS, 1993), (PERELMUTER, 1995).

Muitos dos recentes métodos de Detecção de Faces foram desenvolvidos pela detecção em imagens em tons de cinza, entretanto, a maioria destes pode trabalhar com imagens coloridas (WEI & SETHI, 1999). Esses métodos, primeiramente selecionam prováveis regiões de pele, extraem informações destas regiões e, posteriormente, detectam faces usando padrões faciais.

Em DAI & NAKANO (1996) são isoladas e mantidas as regiões laranja no espaço de cores YIQ como uma região de pele e as demais, eliminadas. Então são utilizadas características de texturas em imagens em escalas de cinza para identificar essas regiões. CHEN *et al.* (1995) adotam um mapa de cores no espaço HSV, para representar possíveis cores de pele. SOBOTTA & PITAS (1996) detectam regiões de pele em imagens usando a matiz e a saturação. MIYAKE *et al.* (1996) utilizam um mapa de cores de pele no espaço CIE Luv. SCHIELE & WAIBEL (1995) trabalham com os componentes verde e vermelho em um grande número de imagens para construir um histograma. O histograma obtido é então empregado para computar a probabilidade de uma combinação particular desses componentes pertencerem à pele. As probabilidades são usadas para classificar os *pixels* da imagem em regiões de pele e não-pele. CROWLEY & BERARD (1997) também adotam esse método para rastrear faces em imagens de vídeo. CAI & GOSHTASBY (1999) empregam o espaço de cores 1976 CIE Lab, descartando a componente luminância, devido a sua variação.

SABER & TEKALP (1998) e WEI & SETHI (1999) usam o espaço de cores YES, onde Y representa a luminância e E e S a crominância. De acordo com SABER & TEKALP (1998), esse espaço proporciona uma melhor representação por

causa de sua eficiência computacional e a ausência de não-singularidades no processo de transformação da cor.

3.2.1.1 Modelo estatístico de cor de pele

FERIS *et al.* (2000) apresentam um modelo estatístico de pele através do espaço de cores RGB. A primeira etapa do método consiste em um treinamento supervisionado a partir de amostras de cor da pele. Um banco de 40 imagens de faces coloridas foi utilizado para realizar o treinamento. Tais imagens foram obtidas de pessoas de diferentes raças, idades e sexo, com variação das condições de iluminação (figura 3.1). No processo de treinamento, uma região de cor da pele é selecionada e sua representação RGB armazenada. A partir dos dados de treinamento, as cores da pele humana formam um agrupamento (*cluster*) no espaço de cores.

FIGURA 3.1 - Amostra de cor da pele.

Embora se espere que essas cores difiram significativamente de pessoa para pessoa, tendo em vista a diversidade de raças existentes, o que realmente ocorre é apenas uma maior diferença de brilho/intensidade e não de cor (YANG *et al.*, 1997). Assim, é possível reduzir a variância do agrupamento de cores da pele através da normalização de intensidade: $a = R / (R + G + B)$, $b = G / (R + G + B)$. As cores (a,b) são conhecidas como cores cromáticas, ou cores “puras”.

De acordo com (YANG *et al.*, 1997) a distribuição de cores da pele no espaço cromático pode ser eficazmente aproximada por um modelo Gaussiano.

Assim, a partir dos dados de treinamento, os parâmetros - média e matriz de covariância - de uma Gaussiana bivariada, podem ser determinados para configurar o modelo de cor da pele.

3.2.1.2 Segmentação da imagem

O procedimento de segmentação consiste em identificar as regiões (*blobs*) de cor da pele em uma imagem de entrada, com base no modelo gerado. Inicialmente, todos os *pixels* da imagem são convertidos para o espaço cromático. Mediante comparação com o modelo Gaussiano, uma nova imagem, em níveis de cinza, é gerada, sendo que a intensidade de cada *pixel* nesta nova imagem representa sua probabilidade de possuir cor da pele (figura 3.2).

FIGURA 3.2 - (a) Imagem de entrada. (b) Imagem obtida a partir do modelo de cor da pele.

Para identificar os *blobs* de cor da pele, a imagem em níveis de cinza é convertida em uma imagem binária, com o ajuste de um valor de limiar (*threshold*) determinado experimentalmente. Em geral, devido a ruídos e distorções na imagem de entrada, o resultado da limiarização pode gerar regiões de cor da pele particionadas (componentes não-conexos), além de permitir a existência de *pixels* de cor da pele isolados na imagem binária. Para solucionar esses problemas é utilizada uma operação morfológica de fechamento seguida da aplicação do filtro da mediana (figura 3.3).

FIGURA 3.3 - (a) Imagem resultante da limiarização. (b) Aplicação do operador de fechamento e filtro da mediana.

A identificação dos limites dos *blobs* de cor da pele é então realizada através do algoritmo de crescimento de regiões e regiões menores que 1% do tamanho da imagem são descartadas. Além disso, são levados em consideração aspectos estruturais da face, eliminando regiões que não apresentam a estrutura de uma face. Isto é feito com base na relação entre a altura e a largura do retângulo que define a região de cor da pele.

É importante salientar que as regiões determinadas podem corresponder a objetos com cor da pele diferentes de uma face. Logo, a verificação da presença ou ausência de uma face em tais regiões faz-se necessária. Isso é feito através de um procedimento simples de correlação, que visa localizar olhos nas regiões de cor da pele (FERIS *et al.*, 2000). Caso existam várias faces na imagem, aquela que estiver mais próxima da câmera será segmentada.

3.2.2 Análise de Componentes Principais

A PCA (*Principal Component Analysis* ou expansão *Karhunen-Loève* – KL) é uma das abordagens mais populares para o problema de reconhecimento de faces (CHELLAPA *et al.*, 1995), (KIRBY & SIROVICH, 1990), (TURK & PENTLAND, 1991), (ROMDHAMI, 1996), (FUKUI & KAWAGUCHI, 1997). Em geral, esse método é utilizado para análise de dados multivariados e tem como objetivo procurar um subconjunto de direções principais num conjunto de dados de qualquer distribuição estatística. A transformação de um conjunto de dados inicial

num espaço de componente principal minimiza a correlação cruzada de uma única amostra de dados (GROß, 1994).

O PCA reduz a dimensão de um conjunto ou espaço para que esta nova base descreva melhor um “modelo” do conjunto. No caso do reconhecimento de faces, o “modelo” é o conjunto de faces a serem treinadas (ROMDHAMI, 1997) (MOGHADDAM, 1995).

KL é um ótimo método linear para reduzir redundância em um conjunto de dados, pois PCA gera um conjunto de eixos ortogonais de projeção conhecida como componente principal (auto-vetores ou *eigenvectors*) dos dados de entrada distribuídos em ordem decrescente de variância. A transformada de KL é um método estatístico bem conhecido para a extração de características e projeção de dados multivariados e tem sido usada amplamente em Reconhecimento de Padrões, Processamento de Sinais, Processamento de Imagens e Análise de dados. Através de KL, pontos em uma entrada espacial de dimensão n são projetados em um espaço de dimensão m , $m \leq n$ (KIRBY, 1990) (VALENTIN, 1994).

3.2.2.1 Uso prático da PCA na procura de características faciais

Os passos do uso da Análise de Componentes Principais para procurar características faciais são mostrados na figura 3.4. No passo inicial, um conjunto de faces, partindo de um banco de um sub-conjunto de faces a serem incluídas, é projetado num sub-espaço definido pelos seus auto-vetores mais importantes. As características dos vetores derivados desse procedimento servem como um treinamento básico para qualquer esquema de classificação supervisionado. Por razões de robustez, usualmente imagens faciais diferentes de uma mesma pessoa são utilizadas. Uma vez “treinada” a classificação, já é possível discriminar qualquer nova e desconhecida imagem por projeção e classificação.

Supondo-se que esse método irá providenciar uma ótima representação das características faciais de diferentes pessoas, um simples classificador euclidiano deve ser suficiente.

FIGURA 3.4 – Esquema geral para reconhecimento de faces usando PCA (GROß, 1994).

3.2.3 Template matching

O método de *template matching* é somente efetivo quando as imagens investigadas possuem o mesmo tamanho, a mesma orientação e a mesma condição de iluminação das imagens na base de dados. Esse método opera de modo a realizar correlações diretas entre segmentos de imagens (CYGANSK, 1996) (BANON, 1997).

O sucesso do *template matching* para o reconhecimento de padrões depende da robustez da biblioteca de padrões. Quanto mais *templates* da mesma face houver na biblioteca, com pequenas variações de posições, melhor será a taxa de acerto no reconhecimento (CYGANSK, 1996) (BRUNELLI, 1993).

A determinação automática de similaridade local entre *templates* de características em todas as locações da área de pesquisa (melhor enquadramento – *best fit*) tradicionalmente tem sido feita utilizando métodos de correlação cruzada ou reconhecimento de padrões (NASA, 1996).

Nesse esquema uma métrica de comparação é utilizada para cada *template* e os valores mais altos obtidos correspondem à face desejada.

O algoritmo de *template matching* usa, via de regra, um produto interno normalizado *pixel-a-pixel* como métrica para comparação. A equação usada para calcular esse valor é mostrada na equação 3.4, onde A e B referem-se às matrizes bi-dimensionais que representam a face e o template, respectivamente.

$$M = \frac{\langle A, M \rangle}{\sqrt{\langle A, A \rangle \langle B, B \rangle}} \quad (3.4)$$

A notação $\langle A, M \rangle$ acima mostra a função do produto interno. O produto interno fornece um método para calcular a projeção de um vetor em outro (ou, nesse caso, uma imagem em outra). Para as duas matrizes A e B de tamanho NxM o produto interno pode ser calculado pela equação 3.5:

$$\begin{aligned}
\langle A, B \rangle &= \sum_{i=0}^{i=N} \sum_{j=0}^{j=M} A_{ij} B_{ij} \\
&= A_{0,0} \cdot B_{0,0} + A_{0,1} \cdot B_{0,1} + \dots A_{N,M} \cdot B_{N,M}
\end{aligned}
\tag{3.5}$$

Para aplicar a equação 3.4, as imagens devem ter o mesmo tamanho, e caso isto não ocorra, devem passar por um pré-processamento que realize a normalização da imagem capturada em relação ao *template* pré-estabelecido.

Apesar de ser um método efetivo em condições ideais, o *template matching* exige características rígidas de similaridade de condições entre captura de imagens, o que pode ser inadequado para muitas aplicações.

A maior desvantagem de *template matching* está em seu custo operacional. No caso do uso de *templates* rígidos, ocorrem falhas ao reconhecer padrões distorcidos ou com grande variação intra-classe.

Uma abordagem usada para detectar e reconhecer o olho, numa imagem frontal, é descrita por (HALLINAM, 1991). O *template* é representado com duas regiões de intensidade uniforme: a primeira é a região da íris e a segunda a esclerótica. Foi construído um protótipo do olho; um conjunto de 107 modelos com variações do mesmo e as mais prováveis localizações errôneas que poderiam ser escolhidas como candidatas ao *template*. BRUNELLI & POGGIO (1993), BEYMER (1994) e KAWAGUCHI *et al.* (2000) também utilizaram templates para localizar os olhos.

3.2.4 Templates deformáveis

Os modelos de *templates* deformáveis podem ser usados quando as deformações não podem ser modeladas diretamente. YUILLE *et al.* (1989), YUILLE *et al.* (1992), CHOW & LI (1993), WANG & TAN (2000) extraem características faciais usando esses tipos de *templates*. A eles são permitidos transladar, rotacionar e deformar para enquadrar-se na melhor representação de formas presentes na imagem. Segundo YUILLE *et al.* (1992), o pré-processamento é feito para obter representações de picos e vales de imagens em níveis de cinza. Filtros morfológicos

são usados para determinar essas representações. Os *templates* para o olho têm 11 parâmetros que consistem dos arcos superiores e inferiores do olho, a íris, os pontos centrais e o ângulo da inclinação do olho. Este *template* é enquadrado na imagem em um sentido de minimização de energia. Determinam-se as funções de energia dos vales, de borda, da imagem, do pico e potencial interno e selecionam-se os coeficientes para cada potencial e emprega-se uma regra atualizada para determinar o conjunto de melhor parâmetro. Nos experimentos, os autores concluíram que a localização inicial do *template* é crítica para determinação da localização exata do olho. Quando o *template* foi deslocado acima da sobrancelha, o algoritmo falhou na distinção entre o olho e a sobrancelha. Outra desvantagem para essa abordagem é sua complexidade computacional. Sempre que aproximações analíticas são feitas na imagem, o sistema precisa ser tolerante a certas discrepâncias entre o *template* e a imagem atual.

3.2.5 Técnicas baseadas em Gabor

A técnica de Gabor mais conhecida é a abordagem local que utiliza *jets*. Um *jet* é um vetor em que cada posição é determinada através de uma transformada de Gabor bidimensional com a janela Gaussiana (modulada por uma exponencial complexa) em um determinado local da imagem, como por exemplo, o centro do nariz, as pupilas, os cantos da boca, etc. Cada campo de um vetor *jet* é determinado pelo cálculo da transformada com uma janela de escala (variância da Gaussiana), orientação e/ou frequência diferente. Cada *jet* é calculado em uma posição diferente de uma imagem (modelo) de treinamento.

Para efetuar a classificação, antes é necessário realizar um processo de combinação com imagens, que se refere ao posicionamento dos *jets* em determinados pontos da imagem. Para isso, um procedimento tenta localizar a posição para cada nó do grafo que, ao mesmo tempo, maximiza a similaridade das características e minimiza o custo da topografia. Esse processo é feito para as imagens de treinamento e de testes, de forma que os pontos importantes da face são localizados e os *jets* são calculados nessas posições. As imagens de teste são classificadas de

acordo com a similaridade com as imagens de treinamento, com base na combinação dos grafos e os *jets* obtidos. Essa abordagem pode apresentar problemas com variações na iluminação e com a imagem do fundo.

Uma outra abordagem é *Wavelet Networks* (ZHANG, 1992). Nessa técnica, uma função é representada por um número fixo de *wavelets*, que correspondem aos nós de uma rede neural. Os coeficientes das *wavelets* são representados pelos pesos sinápticos dessa rede. Mediante um procedimento de otimização, os pesos e parâmetros das *wavelets* são determinados de forma a melhor aproximar a função considerada. É importante salientar que os parâmetros das *wavelets* são obtidos a partir do espaço contínuo de fase, permitindo uma melhor aproximação para um considerado número de *wavelets*. A transformada em *wavelets* pode ser facilmente estendida para o caso 2D (ANTOINE, 1993), com vistas à análise de imagens, por exemplo. Neste caso, porém, além dos parâmetros de dilatação e translação, um parâmetro adicional de rotação é utilizado na transformada.

Para a representação da imagem de uma face usa-se a técnica baseada em **Gabor Wavelet Networks** (GWNs). As GWNs possibilitam que a imagem discreta de uma face seja aproximada por uma combinação linear de *Gabor wavelets*, em que os pesos e parâmetros das *wavelets* são determinados de forma ótima, visando a preservar o máximo de informação da imagem. Os filtros de Gabor asseguram que características geométricas e de textura da face sejam codificadas na representação (KRUGER & SOMMER, 2000). Além disso, o número de *wavelets* da GWN pode ser escolhido de modo a gerar uma representação compacta ou de alta precisão.

As principais características da representação de faces baseada em GWNs estão discriminadas abaixo (FERIS *et al.*, 2000):

- A representação apresenta propriedades de invariância com relação a deformações afins da face e variação homogênea das condições de iluminação. Esta segunda propriedade decorre do fato das *Gabor wavelets* possuírem média zero.
- O grau de precisão da representação depende do número de filtros da GWN, permitindo generalização dos dados de treinamento ou codificação de informação mais precisa da face.

- Os parâmetros das *wavelets* são determinados mediante um processo de otimização de modo a representar o máximo de informação da imagem, considerando um determinado número de filtros.
- Os pesos ótimos de uma GWN estão diretamente relacionados com as respostas dos filtros de Gabor, refletindo a estrutura local da imagem.
- Os filtros de Gabor constituem efetivos detectores de características. Durante o processo de otimização, alguns filtros são atraídos por características locais da imagem, codificando informação geométrica. Os parâmetros de tais filtros tendem a assumir a mesma posição, orientação e escala das bordas da imagem. Outros filtros, porém, codificam informação de textura.
- A representação baseada em GWNs é extremamente compacta. Por exemplo, o espaço ocupado pela representação ilustrada na figura 3.5 (a) consiste em apenas 1040 bytes.

No trabalho de (KRUGER & SOMMER, 2000), essa técnica foi eficazmente utilizada para rastreamento de faces, reconhecimento independente de expressão facial e estimação de pose e no de (FERIS *et al.*, 2000), para tratar de problemas de detecção e rastreamento de faces em seqüências de vídeo.

3.2.5.1 Representação baseada em GWNs

Com a técnica GWN, não somente os pesos, mas também os parâmetros ótimos de cada *wavelet* são determinados a partir do espaço contínuo de fase.

Como exemplo, consideremos a imagem da face ilustrada na figura 3.5(a). Utilizando uma GWN com somente 52 *Gabor Wavelets*, iniciadas na região interna da face, é possível gerar uma representação *wavelet* da face, a qual está ilustrada na figura 3.5(b). É importante mencionar que esta imagem reconstruída, também denominada **Gabor Wavelet Template** (GWT), é uma versão discreta da representação contínua da face. A figura 3.5(c) ilustra a posição das 16 *wavelets* de maior peso, após o procedimento de otimização.

FIGURA 3.5 - (a) Imagem de uma face. (b) Representação obtida pela GWN.
(c) Posição das 16 wavelets de maior peso (FERIS *et al.*, 2000).

A precisão da representação da face é determinada pelo número M de *wavelets* da GWN. A escolha de M deve ser feita com base na aplicação desejada. Por exemplo, para problemas que requerem processamento em tempo real, um número pequeno de *wavelets* pode ser escolhido, gerando-se uma representação compacta e eficiente. Tal representação é adequada para o problema de rastreamento de faces, pois permite eficiência e adaptação a diferentes indivíduos. Por outro lado, se a aplicação exigir o processamento de informação mais precisa da face, então um valor mais alto para M deve ser escolhido.

Esta variabilidade da GWN na representação da informação da face é denominada atenção progressiva. A figura 3.6 ilustra esse conceito, mostrando representações de uma face com diferentes níveis de precisão. É interessante notar que a GWN constitui uma **RBF network** (KOSKO, 1992), permitindo generalização dos dados de treinamento quando um pequeno número de funções base é utilizado.

FIGURA 3.6 - (a) Imagem de uma face. (b) Representações obtidas utilizando GWNs com 32, 52, 100 e 320 *wavelets* (FERIS, 2000).

3.2.6 Método do Triângulo Isósceles

LIN & FAN (2001) desenvolveram uma metodologia para detecção de faces humanas usando relações geométricas do triângulo isósceles. Essa abordagem é subdividida em três passos. No primeiro passo, é lida uma imagem, a qual pode ser colorida ou em níveis de cinza, que será convertida em uma imagem binária. No segundo, rotulam-se todos os componentes 4-conectados na imagem para formarem blocos e encontrar o centro de cada um deles. No terceiro, detectam-se três centros de três diferentes blocos que formarão o triângulo isósceles.

1º passo: Pré-processamento da imagem de entrada para uma imagem binária

A proposta principal dessa etapa é encontrar regiões de face em uma imagem de entrada que possam potencialmente conter faces. A imagem é convertida em uma imagem binária pela limiarização global simples com um limiar único T . Em seguida, é realizada a operação de abertura (erosão primeiro, depois dilatação) para remover ruídos, e então a operação de fechamento (dilatação primeiro, depois erosão) para eliminar buracos.

2º passo: Rotular todos componentes 4-conectados e encontrar o centro de cada bloco

Aqui é usado o “raster scanning” (da esquerda para a direita e de cima para baixo) para obter os componentes 4-conectados, rotulá-los e encontrar o centro de cada bloco.

3º passo: Encontrar quaisquer 3 centros de 3 diferentes blocos que formam um triângulo isósceles

Nesta etapa, procura-se por centros de blocos que formem um triângulo, baseado na regra “a combinação de dois olhos e uma boca”.

Regras para encontrar um triângulo isósceles

FIGURA 3.7: (a) Triângulo isósceles $i j k$. (b) 3 pontos $(i, j \text{ e } k)$ formam um triângulo isósceles

Se um triângulo $i j k$ é um triângulo isósceles como mostra a figura 3.7, então ele deveria possuir a característica de “a distância da linha $i j$ = distância da linha $j k$ ”.

A partir desta observação, LIN & FAN (2001) descobriram que a distância Euclidiana entre os olhos (linha $i k$) é aproximadamente 90% a 110% da distância Euclidiana entre o centro do olho esquerdo/direito e da boca. Devido a imperfeições no resultado da binarização, é dada uma taxa de 25% de desvio. A primeira regra pode ser estabelecida como:

$$|(D(i,j)-D(j,k))| < 0.25 * \max (D(i,j),D(j,k))$$

e a segunda regra como:

$$|(D(i,j)-D(i,k))| < 0.25 * \max (D(i,j),D(j,k))$$

e a terceira:

$$i < j < k$$

$D(i,j)$ é a distância Euclidiana entre os centros dos blocos i (olho direito) e o bloco j (boca), $D(j,k)$ é a distância Euclidiana entre os centros dos blocos j (boca) e o bloco k (olho esquerdo), $D(i,k)$ é a distância Euclidiana entre os centros dos blocos i (olho direito) e o bloco j (olho esquerdo).

Por exemplo, como mostra a figura 3.7a, se 3 pontos (i, j e k) satisfazem as regras, conclui-se que eles formam um triângulo isósceles.

Capítulo 4

A Transformada de Hough

4.1 Introdução

Foi através de uma patente que HOUGH (1962) propôs um método para detecção de padrões complexos, mais precisamente para estudar o comportamento de partículas sub-atômicas através do rastro deixado quando estas passavam através de um campo visual.

A Transformada de Hough (HT) é um método de parametrização onde cada ponto da imagem deve ser representado por uma reta ou por uma senoíde, num espaço de parâmetros.

A HT é uma técnica de grande utilidade em visão computacional. Sua aplicação até meados da década de 80 estava sendo muito lenta devido à complexidade computacional de armazenamento e a dificuldade de um entendimento detalhado de suas propriedades. Contudo, nos últimos anos, muitas pesquisas têm sido realizadas, contribuindo assim para um melhor aproveitamento de seus recursos.

A sua utilização requer que se conheça um pouco sobre a localização das bordas da imagem a fim de que sejam considerados apenas os *pixels* de fronteira. O método de Hough pode ser aplicado para detecção de curvas desde que as mesmas possam ser descritas de forma paramétrica (linha reta, círculos, elipses, etc).

Vários autores têm desenvolvido novos algoritmos, derivados da HT original (DUDA & HART (1972), BALARD (1981), LI *et al.* (1985), ILLINGWORTH & KITTLER (1987), BEN-TZVI & SANDLER (1990), HAN *et al.* (1994)). A implementação em arquiteturas paralelas tem surgido na tentativa de se obter maiores velocidades de processamento (BEN-TZVI *et al.* (1990), ROSENFELD *et al.* (1988)).

4.2 Fundamentos da Transformada de Hough

Pela proposta de Hough, uma linha é mapeada usando os parâmetros Inclinação-Intersecção. Na Figura 4.1a tem-se uma linha reta que conecta uma sequência de pontos expressos pela equação (4.1)

$$y = m x + c \quad (4.1)$$

onde “m” representa a inclinação e “c” a intersecção. Considerando-se 3 pontos pertencentes à mesma reta, cada ponto do espaço imagem é transformado em uma equação de reta, conforme mostra a Tabela 4.1.

TABELA 4.1 – Efeito da HT sobre uma linha reta no espaço m-c

	X	Y	$Y = m * X + c$	$C = -X m + Y$
A	3	4	$4 = m * 3 + c$	$C = -3 m + 4$
B	1,5	1	$1 = m * 1,5 + c$	$C = -1,5 m + 1$
C	0,5	-1	$-1 = m * 0,5 + c$	$C = -0,5 m - 1$

Essas retas são traçadas no espaço de parâmetros, definindo assim, o ponto de intersecção, conforme Figura 4.1b. Todavia, tanto a inclinação como a intersecção, são limitadas e essa parametrização tem a desvantagem de ser sensível à escolha do eixo de coordenadas no plano da figura, o que complica a aplicação da técnica.

FIGURA 4.1 – A Transformada de Hough

a) Espaço de imagem b) Espaço de Parâmetros

Uma solução para esse problema foi proposta por DUDA & HART (1972), que utilizaram coordenadas polares para definir a reta, trabalhando com os parâmetros Ângulo-Raio ao invés de Inclinação-Intersecção. Com isso, os estudos tomaram outro ritmo, proporcionando novos melhoramentos e perspectivas de grandes avanços.

O método de DUDA & HART (1972) envolve um mapeamento de retas do “espaço de imagem” para conjuntos de pontos num “espaço de parâmetros Ângulo-Raio”. No plano da imagem, uma reta parametrizada pode ser representada por dois parâmetros: ρ sendo a distância da normal até a linha da origem da imagem, e θ sendo o ângulo dessa normal como eixo horizontal, como se pode observar na Figura 4.2.

A equação da reta correspondente a essa geometria é:

$$\rho = x \cos \theta + y \sin \theta \quad (4.2)$$

Figura 4.2 – (a) Representação da equação da reta em função de suas coordenadas polares.

(b) Relação entre definição de reta em coordenadas polares.

Os novos parâmetros utilizados para representar o espaço são definidos agora por ρ e θ . Logo, o problema de detectar pontos colineares pode ser convertido no problema de se encontrar curvas concorrentes.

Segundo DUDA & HART (1972), as propriedades para se transformar ponto em curva são:

Propriedade 1 – Um ponto no plano de figura corresponde a uma curva senoidal no plano de parâmetros.

Propriedade 2 – Um ponto no plano de parâmetros corresponde a uma reta no plano de figura.

Propriedade 3 – Pontos pertencentes a mesma linha reta no plano da figura correspondem a curvas através de um ponto no plano de parâmetros.

Propriedade 4 – Pontos pertencentes a mesma curva no plano de parâmetros correspondem a linhas através do mesmo ponto no plano de figura.

Na Figura 4.3 (d) tem-se a propriedade de reflexão, onde os pontos “A”, “B” e “C” se repetem quando a senóide inicia um novo ciclo, entretanto com sinal negativo.

FIGURA 4.3 – Ilustração da HT. (a) Pontos no espaço da imagem. (b) Mapeamento dos pontos no espaço (ρ - θ). (c) Detecção das propriedades de colinearidade. Ponto A denota intersecção de curvas correspondentes aos pontos 3 e 5 no plano $x - y$ e similarmente ponto B corresponde aos pontos colineares 1, 2, 3 e 4. (d) Propriedade reflexão da HT.

De posse das propriedades, é feito o mapeamento das curvas no arranjo acumulador. (A construção do arranjo acumulador bidimensional é definida dentro

de um erro aceitável em ρ e θ formando uma grade). Essa grade pode ser limitada no intervalo $0 \leq \theta < \pi$ e $-R \leq \rho < R$, onde R é o tamanho da célula, desde que pontos fora desse retângulo correspondam a linhas no plano da imagem que não cruzem a célula. Com a limitação de θ , toda linha no plano x-y corresponde a um único ponto no plano $(\rho-\theta)$. Com a equação 4.2, a cada ponto (x,y) do espaço imagem, a curva correspondente é colocada no arranjo incrementando-se de uma unidade o contador em cada célula ao longo da curva.

A análise do arranjo, após os cálculos da equação 4.2, revela células que possuem números elevados de contagens. Esses valores de contagem representam, para cada célula, dentro de um erro de quantização, a quantidade de pontos coincidentes de uma reta. Portanto, o método da HT requer a identificação desses máximos locais. Muitas vezes ocorrem vários picos altos próximos uns dos outros no arranjo. Com isso, existe a necessidade de utilizar um *threshold* para filtrar a quantidade de erros e aproximações e assim poder identificar realmente os parâmetros da imagem.

Os valores de *threshold* estão diretamente relacionados às características de cada imagem.

A Transformada de Hough também pode ser usada para detectar formas circulares. Da mesma forma que na detecção de retas, é possível utilizar um espaço parametrizado para indexar os três parâmetros que definem um círculo: as coordenadas a e b do centro do círculo e o raio. Com esses três parâmetros ficam definidas as posições dos pontos que pertencem ao círculo.

Figura 4.4- Forma circular contida numa imagem.

Para se detectar, por exemplo, a presença de uma forma circular de raio r numa imagem, como o da Figura 4.4, usa-se os N pontos de borda como centros (x_i, y_i) de círculos. Estes círculos são traçados no espaço de parâmetros e, se a forma na imagem for circular, eles interceptarão o centro (a, b) dessa forma, como ilustrado na figura 4.5.

Geralmente o valor exato do raio não é conhecido, e se estabelece então um intervalo de busca para esse parâmetro. Quanto maior esse intervalo, mais valores de raio serão pesquisados. Porém este procedimento provoca um grande aumento do número de valores que precisariam ser calculados e acumulados e também aumenta o intervalo possível para que o processo de detecção se concretize num tempo mínimo, ocupando o menor espaço possível em memória. Por isso, é interessante conhecer com antecedência a faixa de valores possíveis para o raio.

Figura 4.5 – Os círculos paramétricos interceptam-se no centro da imagem.

Quando se faz a variação do valor do raio, o espaço de parâmetros deixa de ser planar e passa a ser espacial. Cada valor de raio será pesquisado num dos planos paramétricos. Dessa forma, ao incrementar o valor do raio forma-se, para cada ponto de borda da imagem, um cone no espaço de parâmetros como ilustrado na Figura 4.6. O ponto de cruzamento de um maior número destes cones mostrará o centro de um provável forma circular na imagem.

Figura 4.6 – Cruzamento entre cones no espaço de parâmetros que determinam a existência de uma forma circular na imagem.

Segundo KIMME *et al.* (1975), o algoritmo de DUDA & HART aborda o método para encontrar formas circulares numa imagem. Nesse método, o raio é limitado ao uso de apenas valores discretos, restringindo assim a faixa e os passos de quantização da forma circular contida na imagem digitalizada. Um exemplo da atuação da grade acumuladora está ilustrado na Figura 4.7.

Podemos visualizar o campo acumulador como planos formados por grades acumuladoras, sendo um plano para cada valor de raio. Círculos de raio r são traçados numa grade sendo os seus centros os pontos de borda da imagem. A cada intersecção destes círculos incrementa-se de 1 (um) o valor acumulado da célula correspondente na grade acumuladora. As células incrementadas um maior número de vezes corresponderão a picos no acumulador nas posições mais prováveis para centros das formas circulares da imagem.

Figura 4.7 – Grade acumuladora quadrangular onde um grande número de círculos passa pela célula correspondente ao centro da forma circular da imagem.

Na detecção de círculos expressa por DUDA & HART, se um círculo é parametrizado pelas coordenadas do centro e o raio (a, b, r) , então estas coordenadas indicam a posição dos pontos de borda (x, y) através da equação 4.3

$$(x - a)^2 + (y - b)^2 = r^2 \quad (4.3)$$

Esta equação também indica que um dado ponto de borda (x_i, y_i) deve estar incluído em algum círculo cujos parâmetros estão contidos na superfície de um certo cone circular no espaço de parâmetros (a, b, r) . Se no cone houver vários pontos de borda interceptando-se no mesmo ponto, então todos os pontos de borda da imagem que contribuem para este ponto estarão no círculo definido por esses parâmetros. A figura 4.8 ilustra a situação de vários planos formados por grades acumuladoras.

Figura 4.8 – Planos constituídos de grades acumuladoras quadrangulares onde um grande número de círculos passa na célula correspondente ao centro da forma circular na imagem.

DAVIES (1992) fez um estudo dos picos obtidos no acumulador da Transformada de Hough. Para fazer o estudo, ele usou um acumulador para detecção de círculos e concluiu que a complexidade do acumulador depende da qualidade da imagem que está sendo verificada. Além disso, a acuidade com a qual o objeto pode ser localizado depende fundamentalmente da forma do pico.

KIMME *et al.* (1975) descreveram um método para detectar formas circulares através da direção do gradiente de cada ponto de borda da imagem. Quando as direções do gradiente desses pontos apontam para uma mesma célula acumuladora, esta célula é então incrementada. A posição do centro é escolhida passando-se um *threshold* proporcional ao raio, de forma que a razão entre o valor do *threshold* e do raio seja sempre constante. O uso da direção do gradiente, segundo eles, é justificado pela sensível diminuição do tempo de processamento.

Essa idéia também é usada por YUEN *et al.* (1990) para detectar o centro de uma possível forma circular na imagem. O centro do círculo paramétrico deve estar contido na linha que passa através do ponto de borda e ao longo da direção do gradiente. Como resultado, em vez de incrementar todo o cone circular, apenas segmentos do cone precisam ser incrementados. Este procedimento foi denominado de Transformada de Hough Padrão (SHT).

O acumulador deve ser incrementado em 3D (a,b,r) , podendo ser eficientemente indexado usando as direções do gradiente que irão definir o centro (a,b) e a equação do círculo a fim de se calcular o parâmetro r . Uma escolha natural é ter cada célula de grade acumuladora representada por um ponto no espaço de parâmetros. Desta forma, para procurar por uma forma circular numa imagem de dimensão N^2 , cujo raio é menor ou igual a N , é necessário um acumulador 3D de tamanho N^3 .

Uma complementação a se fazer a respeito do uso da direção do gradiente é que, se por um lado permite uma redução de memória requerida, por outro, se a imagem for muito ruidosa ou complexa, este método tende a ser ineficaz, detectando falsas formas circulares.

ILLINGWORTH & KITTLER (1988) fizeram um levantamento sobre a Transformada de Hough e concluíram que esse método possui grande capacidade de detecção de formas, mesmo sob condições de ruído, porém a necessidade de mais memória e tempo de processamento cresce com o aumento da dimensionalidade do acumulador.

Na tentativa de resolver o problema da dimensionalidade do acumulador, CASASANT & KRISHNAPURAM (1987) assumiram que um objeto curvo pode ser considerado com uma composição de pequenos segmentos de retas tangenciais à

borda desse objeto. Cada um desses segmentos corresponde a um ponto (ρ, θ) no espaço paramétrico. Deste modo, irá existir um grande número de pequenos segmentos de retas e estes irão gerar pequenos picos no espaço paramétrico.

Picos menores ainda, geralmente de altura igual a dois, irão ser gerados. Esses picos são provenientes de retas que possuem apenas dois pontos na imagem. Essas retas geralmente não são tangenciais à curva e, portanto, não devem contribuir para a detecção da curva na imagem e deverão ser retirados através de um *threshold*.

Depois de retirados os picos muito pequenos, os pontos (ρ, θ) incrementados maior número de vezes, irão cada um caracterizar a coordenada polar de um ponto na imagem. Assim, através de uma Transformada inversa de Hough será possível detectar a curvatura.

Este método pode ser usado para detectar quaisquer objetos curvos na imagem, tais com: objetos circulares, elípticos e parabólicos, mesmo que estejam parcialmente ocultos ou seccionados. Uma grande vantagem deste método é a redução de memória requerida, já que trata objetos curvos (que na HT convencional requerem um acumulador multi-dimensional) como segmentos de retas, que requerem acumuladores bidimensionais. Porém, este método possui um grande inconveniente na escolha do *threshold*, pois para uma seleção razoável é necessário conhecer a resolução da entrada e a do espaço paramétrico. Em imagens muito complexas, este método deverá falhar, pois os picos referentes à imagem real, por serem pequenos, tendem a ter a mesma amplitude dos picos referentes às retas que não são tangenciais à curva, sendo assim impossível separá-los utilizando um *threshold*.

HANAHARA & HIYANE (1990), baseados na analogia da HT com o Princípio de Huygens's, apresentaram um algoritmo de detecção de formas circulares que resolve numericamente uma equação de onda bidimensional usando apenas operações com bases adjacentes, que são, Laplaciano, adição de imagem e multiplicação de constantes com o conteúdo da imagem. Pelo fato de não usar extração de borda, o algoritmo detecta formas circulares em imagens de baixo contraste.

Este método usa o Princípio de Huygens's – uma regra matemática para descrever o comportamento da luz – o qual estabelece que a superposição de todas as

ondas esféricas elementares com mesmo raio, emitidos a partir de posições puntiformes numa frente de onda, forma um estado de onda futuro.

A HT é muito sensível à complexidade da imagem e deve falhar nas imagens de baixo contraste. O problema não é propriamente da Transformada, mas da extração da borda (usada para simplificar a imagem e diminuir a quantidade de processamento), e mais especificamente da escolha do *threshold*.

O conceito básico deste método vem das ondas na água, com a altura das ondas correspondendo à escala de cinza. Distúrbios simultâneos na circunferência de um círculo irão se propagar na direção do centro, onde formarão um pico. Dessa maneira, simulando-se a propagação de onda, identificam-se picos que podem ser usados para detectar formas circulares.

Algoritmos que se valem da Transformada de Hough e do Princípio de Huygens's desenham ambos círculos ou esferas com o mesmo raio e usam o princípio de superposição. Na Transformada de Hough, o princípio da superposição aparece durante o armazenamento, e no Princípio de Huygens's essa propriedade é inerente à equação da onda.

A desvantagem desse método é o fato de possuir um custo computacional maior do que o método baseado na HT. Enquanto neste último a quantidade de processamento é proporcional a $R^2 N_{\text{borda}}$, no método proposto por HANAHARA & HIYANE (1991) essa proporcionalidade fica em RN^2 , sendo R o número de raios pesquisados, N uma das dimensões de uma imagem NxN e N_{borda} o número de pontos de borda da imagem.

Um dos problemas do SHT de YUEN *et al.* (1990) é a grande quantidade de dados a serem armazenados se o intervalo de valores para o raio for muito grande. GERIG & KLEIN (1986) então reordenaram o cálculo da HT para refazer o acumulador 3D, de dimensão N^3 , com três acumuladores 2D, ordenados em tamanho N^2 . Este método é chamado de Transformada de Hough de Gerig (GHT). Ao contrário do SHT, onde a determinação do pico é feita usando todos os parâmetros, a GHT realiza a HT em estágios.

Uma deficiência do método de GERIG & KLEIN (1986) é que devido ao fato dos picos serem gravados em posições espaciais bidimensionais, os valores serão reescritos se dois círculos paramétricos interceptarem uma posição comum. Por

essa razão, numa implementação simples não é possível acumular todos os detalhes de círculos concêntricos.

A eficiência desse método pode aumentar se neste for incorporado a direção do gradiente, ao invés de se acumular todo o círculo paramétrico. Essa modificação da GHT é denominada de Transformada de Hough de Gerig com Gradiente (GHTG). E se a direção do gradiente for usada, então apenas a linha do cone que cai sobre a direção do gradiente de um ponto de borda será incrementada.

Se a direção do gradiente dos pontos de borda é disponível, uma outra forma de reduzir o armazenamento e demanda computacional na procura pela forma circular é decompor o problema em dois estágios como mencionado em ILLINGWORTH & KITTLER (1988). O método consiste num acumulador 2D para encontrar o centro da forma circular, seguido de um HT de 1D para determinar o raio. Esse método é por eles chamado de Transformada de Hough 2-1 (2-1HT). O centro da forma circular deve estar ao longo da posição do gradiente de cada ponto de borda. Então o ponto de intersecção comum desses gradientes identifica seu centro.

O primeiro estágio é usado para acumular os valores de HT indexando com os valores do centro (a,b) num espaço 2D e o segundo estágio usa os parâmetros do centro do primeiro estágio e pela equação do círculo, constrói e analisa um histograma de possíveis valores de raio.

Um histograma do raio é feito para cada candidato a centro do círculo, e picos no histograma do raio indicam evidência de círculos. O processo é ilustrado nas figuras 4.9 e 4.10.

Um grande problema do método 2-1HT é o fato do acumulador 2D usado para encontrar o centro ser mais difícil de se analisar que o acumulador 3D, o qual vem a ser uma projeção completa daquele espaço. Em adição, um falso pico ou inacuidade na estimativa do parâmetro no primeiro estágio irá afetar a performance do segundo estágio.

LI *et al.* (1986) sugeriram um método denominado de Transformada de Hough Rápida (FHT) que usa uma estrutura multi-dimensional para simultaneamente acumular e detectar picos, necessitando de menos memória que a SHT para seu armazenamento.

O parâmetro original de volume é considerado com um hipercubo simples, se o número de ponto da imagem que incrementa os parâmetros de um certo hipercubo for maior do que um valor limite. Então assume-se que no espaço de parâmetros correspondente existe um pico, e o volume do hipercubo é investigado com maior detalhe pela decomposição do hipercubo em sub-hipercubos.

Figura 4.9 – O centro é localizado através de um espaço acumulador bidimensional

Figura 4.10 – O valor do raio é encontrado através de um histograma obtido a partir da medida entre os pontos de borda da imagem e o valor do centro.

Este é um processo iterativo de acumulação, análise e divisão contínua até que seja atingido um limite de resolução dos parâmetros, ou o hipercubo falha e seus incrementos ficam abaixo do limite que designe um pico.

Uma modificação apropriada para o método de Li *et al.* (1986) denominada de Transformada de Hough Rápida Modificada (MFHT) utiliza a direção do gradiente de cada ponto da imagem.

ILLINGWORTH & KITTLER (1987) introduziram a Transformada de Hough Adaptativa (AHT), que usa um pequeno acumulador ordenado ($9 \times 9 \times 9$). Realiza uma análise grosseira de cada parâmetro assim como na FHT de LI *et al.* (1986), e iterativamente focaliza os parâmetros das regiões de maiores densidades do acumulador. A escolha da resolução de cada parâmetro para próxima interação é independente entre si, e é feita de maneira dinâmica e depende da distribuição do número de incrementos no acumulador.

Isto supera o principal problema associado com a parametrização polinomial da elipse, que é o de grande armazenamento de dados devido às grandes larguras de faixa dos parâmetros.

Por outro lado, problemas com a AHT incluem sua complexa implementação. Além disso, assim como no método de Li *et al.* (1986), células incrementadas pelas regiões de maiores densidades de pontos de borda tendem a ser mais incrementadas gerando assim falsos picos no acumulador. Desse modo, o método tende a falhar quando a imagem a ser analisada for complexa.

Também com o objetivo de diminuir a dimensão do acumulador e aumentar a velocidade de processamento, CHAN & SIU (1991) apresentaram um método que usa apenas um par de acumuladores bidimensionais ordenados. Este método é capaz de discriminar múltiplas (incluindo concêntricas) formas circulares em imagens complexas.

As duas dimensões do primeiro acumulador representam as coordenadas **a** e **b** dos centros das formas circulares e as do segundo acumulador representam a coordenada **a** do centro e o raio. Cada ponto de borda irá ter uma reta correspondente neste último espaço acumulador obtida a partir da equação 4.3. A intersecção entre retas determinará o valor do raio para um determinado valor do parâmetro **a**.

A Figura 4.11 mostra que a capacidade de encontrar formas circulares concêntricas através desse método é mantida, enquanto que os acumuladores utilizados são bidimensionais reduzindo a necessidade de memória.

Figura 4.11- a) Imagem contendo formas circulares concêntricas; b) Espaço acumulador dos parâmetros a e b ; c) Espaço acumulador dos parâmetros a e r .

Por se tratar de um método de particionamento do acumulador, este método apresenta os mesmos problemas relatados para o método 2-1HT.

YUEN *et al.* (1990) compararam vários métodos de detecção de formas circulares. Os métodos considerados são: a Transformada de Hough Padrão (SHT), a Transformada de Hough Rápida Modificada (MFHT) de Li *et al.* (1986), o método que executa Hough em dois estágios (2-1HT) mencionado em ILLINGWORTH & KITTLER (1987), o método que utiliza apenas acumuladores bidimensionais (GHT), e por fim, aquele que, além disso, utiliza a direção do gradiente (GHTG) descritos por GERIG & KLEIN (1986).

A fim de comparar os vários métodos, YUEN *et al.* (1990) usaram uma imagem contendo muitas formas circulares de diversos raios, com pouco ruído.

O gráfico da razão entre o número de formas circulares detectadas e o número de formas circulares presentes na imagem em função dos métodos utilizados está ilustrado na Figura 4.12.

Figura 4.12 – Gráfico de desempenho de cada método em relação ao número de formas circulares detectadas.

O resultado em termos de tempo de processamento e quantidade de armazenamento necessário do estudo comparativo está no gráfico ilustrado na Figura 4.13. Pode-se perceber por esse gráfico que a SHT requer muito espaço para armazenamento de dados enquanto a GHT necessita de um longo tempo de processamento. Tanto a 2-1HT quanto a GHTG consegue ter um bom compromisso entre o armazenamento e o tempo computacional. A performance da MFHT depende da complexidade espacial da imagem que está sendo analisada. A comparação entre a GHT e a GHTG ilustra a vantagem computacional do uso da direção do gradiente quando a imagem é isenta de ruído.

Nessas condições a GHTG aparece como o melhor método, pois requer pequeno armazenamento, é rápido, preciso e tem boa relação entre as formas detectadas e as totais, na imagem. Seu problema é a incapacidade de detectar círculos concêntricos.

A SHT requer enorme armazenamento. A 2-1HT é razoavelmente boa e requer pequeno armazenamento, mas tem problemas na detecção de formas circulares pequenas. A Transformada de Houhg de Gerig é considerada ineficiente pela falta do uso da direção do gradiente e a MFHT é comparativamente grande e dependente dos detalhes da estrutura da imagem.

Figura 4.13 – Gráfico comparativo do tempo gasto para armazenamento dos diversos métodos.

4.3 Principais modelos de Transformada de Hough

4.3.1 Transformada de Hough Generalizada

A Transformada de Hough Generalizada (GHT) é uma variação da HT original. BALLARD (1981) utilizou um mapeamento entre o espaço da imagem e o espaço da HT. A imagem inicialmente é tratada por operadores de borda. Os operadores mais comuns são de dois tipos: o operador Gradiente (ex: PREWITT apud BALLARD (1981)) e o operador de HUECKEL (1973), os quais transformam os níveis de cinza da imagem em uma rampa e um degrau, respectivamente. Com isso, cada ponto na imagem não mais contém informações do nível de cinza mas, ao contrário, cada ponto contém uma magnitude e direção representando a orientação da mudança do nível de cinza local.

No algoritmo, o mapeamento é definido a partir da orientação das bordas em relação a um ponto de referência. O ponto de referência é um ponto arbitrário, sendo considerado como a origem do sistema de coordenadas locais para o formato. O grande segredo para a identificação de formatos arbitrários da GHT é o uso da informação direcional. Tal informação, além de permitir uma identificação mais rápida do formato, também melhora sua precisão. Por exemplo, na detecção de

círculos, sem a informação direcional, qualquer grupo significativo de pontos de borda teria grande probabilidade de ser detectado.

No arranjo acumulador, um problema que surge é na detecção dos pontos máximos. Muitas vezes, pontos na vizinhança do ponto ideal são incrementados, ao invés do próprio ponto. Um meio de diminuir esse problema é utilizar um modelo de erro formal na etapa de incrementação como pode ser visto na Figura 4.14. Esse modelo pode especificar um conjunto de pontos próximos ao invés de um único ponto. SHAPIRO (1978a) (1978b) (1975) fez extensivo trabalho neste assunto.

Trabalhos realizados anteriormente utilizaram imagens de bordas binárias para detectar curvas analíticas em DUDA & HART (1972) e não analíticas em MERLIN & FABER (1975). Uma outra proposta feita por BALLARD (1981) foi a detecção de curvas analíticas em imagens em níveis de cinza; já O'GORMAN & CLOWES (1976) estudaram especialmente linhas retas; SHAPIRO (1978a) (1978b) elaborou várias propriedades para detecção de círculos e parábolas em imagens ruidosas.

Figura 4.14 – Usando convolução para compensação de erros na GHT.

Em imagens arbitrárias, para cada formato de imagem a GHT cria uma tabela, a tabela "R" onde serão colocadas as distâncias de pontos da superfície ao ponto de referência, conforme mostra a Figura 4.15. Dado o limite do formato, a tabela "R" pode ser facilmente construída examinando os pontos limites do formato e requer um número de operações proporcional ao número de pontos do limite. Segue sua forma:

- escolhe-se um ponto de referência y para o formato;
- para cada ponto limite x_n da borda computar a direção do gradiente ($\phi(x)$) e o valor de $r = y - x_n$ (distância da borda da imagem ao ponto de referência);
- com os valores de r indexados aos valores de ϕ constrói-se a tabela “R”.

FIGURA 4.15 – Geometria da Transformada de Hough Generalizada.

Uma das propriedades importantes da GHT é o fato de se poder fazer mudanças de escala, rotação, reversões da base da figura, e translação do ponto de referência de r . Todas essas propriedades podem ser computadas pelas modificações diretas para a tabela “R”.

4.3.2 Transformada de Hough Rápida

Na HT convencional, vista em DUDA & HART (1972), cada ponto do traço no espaço da imagem gera “votos” para um conjunto de pontos no espaço de parâmetros. Já LI *et al.* (1986) propuseram um novo enfoque à HT, usando os princípios de formulação do hiperplano para o conjunto de pontos do espaço de parâmetros. Esses hiperplanos são divididos gradativamente por um algoritmo em hipercubos de baixa e alta resolução. Em seguida é realizada uma contagem nos hiperplanos e um valor de “threshold” é determinado, inicialmente, para limitar a criação de subdivisões de hipercubos. A decisão se um hipercubo recebe um voto de um hiperplano dependerá do hiperplano cruzar o hipercubo. Esse algoritmo foi chamado de Transformada de Hough Rápida (FHT). Esse enfoque hierárquico leva a

uma redução significativa da complexidade computacional e o requerimento de armazenamento simultâneo.

Devido ao problema da formulação do hiperplano e à representação *K-tree* do espaço de parâmetros, o teste de intersecção pode ser feito incrementalmente, sem multiplicação. Isso é importante para uma futura implementação de hardware e para extensões do método para manejar superfícies de ordem maior no espaço de parâmetro.

FIGURA 4.16 – Exemplo de um processo de FHT

Através da figura 4.16 é possível ter uma visão geral do processo da FHT. Essa figura é um hiperplano bidimensional formado por oito retas.

Inicialmente, o hiperplano é dividido em quatro hipercubos com endereços (0,0), (0,1), (1,0), (1,1). Escolhendo um *thresholding* - $T=7$ observa-se pelos endereços de “K-Tree”, mostrados na Tabela 4.2, que o hipercubo (1,1) obteve um ponto formado por 8 cruzamentos. Em seguida, o hipercubo (1,1) é dividido em 4 hipercubos (de tamanho $\frac{1}{4} \times \frac{1}{4}$). Nota-se que o hipercubo (11,11) depois de um *thresholding* encontrou o maior cruzamento. Uma nova divisão é feita de tamanho $\frac{1}{8} \times \frac{1}{8}$ e encontrando o valor 8. Tal processo de divisão pode ser repetido até a desejada resolução ser alcançada.

TABELA 4.2 – Valores referentes ao exemplo do processo de FHT

Voto	Endereço K-Tree			
Tamanho $\frac{1}{2} \times \frac{1}{2}$	(0,0)	(0,1)	(1,0)	(1,1)
	3	3	4	8
Tamanho $\frac{1}{4} \times \frac{1}{4}$	(10,10)	(10,11)	(11,10)	(11,11)
	3	5	6	8
Tamanho $\frac{1}{8} \times \frac{1}{8}$	(110,110)	(110,111)	(111,110)	(111,111)
	8	4	4	3

O armazenamento requerido pelo exemplo dado através da Figura 4.7 são 4 registradores de cruzamentos em cada nível de divisão. Uma implementação direta do método de Hough (no qual o espaço de parâmetro é uniformemente dividido e cada hipercubo é testado contra cada hiperplano), poderia ter requerido 64 registradores de “voto” para a escolhida dimensão e quantização. O algoritmo da FHT executa o teste para 13 hipercubos ($1 + 4 + 4 + 4$) de diferentes resoluções, enquanto que um enfoque direto executa o teste para 64 hipercubos.

Alguns problemas surgem em imagens complexas. O valor de *threshold* pode ser difícil de se determinar e, se esse for colocado muito baixo, a eficiência do método pode deteriorar-se, pois o algoritmo irá explorar muitos quadrantes. Outro problema é relacionado ao método de teste de intersecção incremental. Isso requer traço da imagem ao seu centro. Se existem muitos traços na imagem, isso pode representar uma grande sobrecarga.

O algoritmo FHT é altamente paralelo. O processamento para os hipercubos é independente uns dos outros. Além disso, o teste de intersecção para um hiperplano não depende de outros hiperplanos. Quando a implementação paralela é considerada, surge o problema de sincronização.

A sincronização envolve manter a ordem parcial da escolha política de controle para a FHT. A política de classificação de contagem de “voto” é mais difícil de manter.

4.3.3. Transformada de Hough Adaptativa

ILLINGWORTH & KITTLER (1987) desenvolveram a Transformada de Hough Adaptativa (AHT). A idéia principal é uma estratégia iterativa de procura “Grosseira para Fina” para detectar linhas e círculos em espaços de parâmetros 2D e 3D.

Na transformada padrão, quando o arranjo acumulador torna-se muito grande, começam a surgir aspectos indesejáveis, o que requer um espaço de armazenamento muito grande. Se K parâmetros têm de ser determinados e cada faixa de parâmetro é dividida em L intervalos, o armazenamento requerido é Lk elementos. A AHT utiliza, inicialmente, um arranjo acumulador pequeno e de tamanho fixo. Na primeira etapa, cada parâmetro é resolvido apenas muito grosseiramente. Tem-se assim uma idéia básica da imagem.

A adequação da AHT para detectar formatos de linha e círculo depende da identificação correta das regiões de alta densidade que produzem no espaço de parâmetros, mesmo quando o espaço é grosseiramente quantizado.

No algoritmo adaptativo, uma das primeiras decisões que deve ser tomada é definir o tamanho do arranjo acumulador, que geralmente é pequeno.

ILLINGWORTH & KITTLER (1987) escolheram um arranjo 9x9, o qual atendeu às expectativas. Os dados acumuladores são inteligentemente usados para redefinir a faixa de parâmetros, de modo que áreas interessadas possam ser investigadas em maiores detalhes.

Um importante aspecto do algoritmo é sua habilidade para redefinir parâmetros, independente de cada iteração. Os critérios usados para identificar o melhor pico são o valor resultante da divisão do número de contagens pelo número de células acumuladoras. Tomando como referência a posição do melhor pico, calcula-se em torno deste quantas iterações forem necessárias para identificar a imagem.

Na Figura 4.17a foram necessárias sete iterações para determinar os parâmetros da linha. A Figura 4.17b mostra as retas que são geradas de parâmetros da AHT. Na Figura 4.17c tem-se o arranjo acumulador. Os três pontos hachurados representam o possível ponto de intersecção da reta.

Figura 4.17 – Detecção de segmentos lineares usando a AHT. (a) Imagem. (b) Espaço de parâmetros. (limites $m = -1$ a 1 , $c = -256$ a 256). (c) Acumulador grosseiramente quantizado.

Na Figura 4.18a foram necessárias 4 iterações para determinar o centro do círculo. No arranjo acumulador, na Figura 4.18c, observa-se apenas um ponto hachurado, isso facilita a interpretação. A célula sombreada é o pico selecionado pela limitação do arranjo.

O enfoque dos métodos de LI *et al.* (1986) é similar à AHT, à medida em que o espaço de parâmetros é investigado em várias resoluções, mas com um maior detalhamento nas regiões onde ocorre uma alta densidade de pontos. Entretanto, esses métodos diferem da AHT quanto ao grau de liberdade permitida na redefinição dos limites do parâmetro, isto é, o grau de flexibilidade na colocação e escolha do formato da janela que define a faixa de parâmetros em estudo.

Figura 4.18 – Detecção do centro do círculo usando a AHT. (a) Imagem. (b) Espaço de parâmetros (limites $a = 0$ a 128, $b = 0$ a 128). (c) Acumulador grosseiramente quantizado.

Quando a imagem torna-se muito complexa, os efeitos da quantização tornam sua interpretação mais difícil. Nesse caso, é necessário adotar estratégias com um certo grau de complexidade, envolvendo combinações de deslocamento, expandindo e reduzindo a faixa de parâmetros do atual interesse. Muitas vezes o método tem sido desviado para resoluções grosseiras, onde padrões de votos estendidos de dois ou mais objetos podem sobrepor-se e produzir um pico falso, cuja posição não é relacionada aos parâmetros de qualquer traço real na imagem.

Capítulo 5

Localização de olhos através da Transformada de Hough

Elemento fundamental no processo de comunicação visual, a face é o principal atributo fisiológico através do qual as pessoas são reconhecidas visualmente umas pelas outras. Este fato sugere que o desenvolvimento de sistemas automáticos de detecção, rastreamento e reconhecimento de faces venha a oferecer facilidades em aplicações como autenticação de usuários, monitoramento de ambientes, controle de acesso e vigilância. O Reconhecimento de Faces constitui um foco recente e bastante importante de pesquisa na área de Visão Computacional.

Dentre as diversas características faciais disponíveis, os olhos representam o mais importante papel em normalização de faces e facilita, posteriormente, a localização de outros elementos faciais (BRUNELLI & POGGIO, 1993), (LAM & YAN, 1998), (HAN *et al.*, 2000), (KAWAGUCHI *et al.*, 2000), (SIROHEY & ROSENFELD, 2001), (HUANG *et al.*, 1998), (TOLBA *et al.*, 2000), (LAM & LI, 1998), (YUILLE *et al.*, 1989), (SMERALDI *et al.*, 2000), (WANG & TAN, 2000), (FUKUI & YAMAGUCHI, 1997).

Neste trabalho são estudadas algumas técnicas para a detecção de olhos em faces humanas, tendo sido implementadas as etapas do fluxograma da Figura 5.1. A fim de comprovar a robustez e eficiência do método escolhido, foram feitos testes comparativos com outro método: geometria do triângulo isósceles, descrito no capítulo 4.

FIGURA 5.1 : Fluxograma do algoritmo desenvolvido

5.1 Metodologia

Foram utilizadas 100 imagens de um banco de imagens padrão para reconhecimento de faces, com resolução espacial de 200×200 *pixels*. Essas imagens, apesar de serem frontais, incluem algumas variações, tais como de expressão, olhos fechados e rotação. A figura 5.2 ilustra algumas das imagens adquiridas e em uso durante os testes realizados neste trabalho.

FIGURA 5.2 – Imagem do banco de imagens obtido da Internet

5.1.1 Pré-processamento

A imagem resultante do processo de aquisição pode apresentar diversas imperfeições, tais como: presença de *pixels* ruidosos, contraste e/ou brilho inadequado, traços interrompidos ou indevidamente conectados, etc. A função da etapa de pré-processamento é aprimorar a qualidade da imagem com aplicação de filtros visando a eliminação de ruídos provenientes da aquisição. Para isso foi utilizado um filtro passa-baixa, o qual, no domínio do espaço, pressupõe a aplicação

de uma máscara como a que aparece na figura 5.3. Este tipo de filtro é normalmente utilizado para suavizar imagens, principalmente para eliminar (atenuar) ruídos. (GONZALES, 1993).

$$\frac{1}{3} \begin{array}{|c|c|c|} \hline 0 & 1 & 0 \\ \hline 0 & 1 & 0 \\ \hline 0 & 1 & 0 \\ \hline \end{array}$$

FIGURA 5.3 – Exemplo de filtro passa-baixa no domínio do espaço.

5.1.2 Segmentação

A tarefa básica da etapa de segmentação é dividir uma imagem em suas unidades significativas, ou seja, nos objetos que a compõem. Esta tarefa, apesar de simples de descrever, é uma das mais difíceis de implementar. Para realizar a segmentação, foi utilizado o operador de Sobel, descrito a seguir.

5.1.2.1 Operador de Sobel

O operador de Sobel é aplicado na imagem através das máscaras que aparecem na figura 5.4 (GONZALEZ, 1993).

-1	0	1	1	2	1
-2	0	2	0	0	0
-1	0	1	-1	-2	-1
G_x			G_y		

FIGURA 5.4 – Máscaras utilizadas no detector de Sobel

As aplicações dos dois operadores G_x e G_y resultam nos gradientes da borda na direção x e em y, respectivamente e, por meio destes, é possível a obtenção da magnitude e da direção (ângulo) da borda em cada *pixel*, como mostra a Figura 5.5.

FIGURA 5.5 – Determinação da Magnitude dos gradientes e também do ângulo.

A magnitude e a direção das bordas são dadas por pelas equações (5.1) e (5.2).

$$\text{Magnitude}_{G_x, G_y} = \sqrt{G_x^2 + G_y^2} \quad (5.1)$$

$$\alpha = \arctan\left(\frac{G_y}{G_x}\right) \quad (5.2)$$

É importante notar que, pela forma como as máscaras são construídas (baseadas no operador diferencial), as mesmas apresentam uma resposta nula em regiões homogêneas (onde a derivada é nula). Magnitudes de pequeno valor indicam que o *pixel* em estudo não pertence a uma borda bem definida (área mais homogênea).

5.1.2.2 Limiarização (*THRESHOLDING*)

Um método eficiente de limiarização deve ser capaz de fornecer automaticamente um valor (limiar), para o qual, todos os *pixels* com valor de brilho inferior a este limiar devem ser eliminados (magnitude igual a zero, e sem direção).

Existe um conjunto muito grande de técnicas para a obtenção de um valor limiar, e observa-se que cada uma delas utiliza um critério que considera importante para obter o valor limiar, como limiarização global simples, P-Tile, Método de Pun,

Método de Kapur, Sahoo e Wong, Método de Johannsen e Bille, Método do Triângulo, (GONZALEZ, 1993), (PARKER, 1996), (SAHOO *et al.*, 1988) e (HUSSAIN, 1991). Neste trabalho será utilizado o método de limiarização global simples por quatro razões: as imagens do banco possuem um baixo contraste, fácil implementação do método, obtenção de resultados satisfatórios e o fato da técnica escolhida para comparação com a HT, método do triângulo isósceles, também utilizar este tipo de limiarização.

Limiarização global simples – Esta técnica de limiarização é a do particionamento do histograma da imagem por um limiar único T . A segmentação é então efetuada, varrendo-se a imagem, *pixel a pixel*, e rotulando-se cada *pixel* como sendo do objeto ou do fundo, dependendo se o nível de cinza daquele *pixel* for maior ou menor que T . Uma imagem limiarizada $g(x,y)$ é definida como:

$$g(x,y) = \begin{cases} 1 & \text{se } f(x,y) > T \\ 0 & \text{se } f(x,y) \leq T \end{cases} \quad (5.3)$$

Com isso, *pixels* rotulados como 1 (ou qualquer outro nível de cinza conveniente) correspondem aos objetos, enquanto que aqueles rotulados com 0 correspondem ao fundo. (GONZALEZ, 1993). A Figura 5.6 apresenta as imagens resultantes do processo de suavização e de segmentação.

FIGURA 5.6 – (a) Imagem de entrada; (b) resultado do filtro da média 3x3; (c) resultado de Sobel

5.2 Transformada de Hough para objetos circulares

DUDA & HART (1972), a partir da utilização da transformada de Hough para retas, adaptou-a para circunferências. A transformada de Hough é aplicada a uma circunferência dada por uma equação paramétrica. A equação de uma circunferência é

$$(x-a)^2 + (y-b)^2 = c^2 \quad (5.4)$$

onde “a” e “b” são coordenadas do centro da circunferência e “c” o seu raio de curvatura.

Tomando-se todos os *pixels* do plano imagem e aplicando-se a equação 5.5, tem-se que cada *pixel* do plano imagem gerará uma circunferência no plano de parâmetros (a,b) que será armazenada no arranjo acumulador (a, b, r). O cruzamento das circunferências traçadas no plano de parâmetros (a,b) define as coordenadas do centro de uma circunferência e o valor acumulado nesta célula define quantos *pixels* pertencem à circunferência.

$$(x_i-a)^2 + (y_i-b)^2 = c^2 \quad (5.5)$$

Em LOW (1991) encontra-se como detectar circunferências com a Transformada de Hough escolhendo-se a representação paramétrica da mesma. Assim, mudando-se do plano de imagem para o plano de parâmetros tem-se:

$$(a - X_i)^2 + (b - Y_i)^2 = c^2 \quad (5.6)$$

A aplicação da transformada de Hough consiste em mudar a imagem do plano cartesiano para o plano de parâmetros. Isto significa que cada *pixel* do plano cartesiano (imagem) será convertido em circunferência no plano de parâmetros.

FIGURA 5.7 – Imagem com uma circunferência.

Dessa forma, dada uma imagem com cinco *pixels*, como a apresentada na figura 5.7, para se descobrir se existem *pixels* que possam compor a imagem de uma circunferência de raio $= 1/\sqrt{2}$, monta-se um arranjo acumulador com o plano de parâmetros, conforme ilustra a figura 5.8

FIGURA 5.8 – Plano de parâmetros “a” e “b” para circunferências.

O arranjo acumulador é o local onde serão armazenadas as circunferências traçadas com centro nos “*pixels* da imagem” (plano cartesiano). Em princípio, pode ser definido com as dimensões iguais às da imagem no plano cartesiano, com tamanho “a” x “b” onde o tamanho de “a” é igual ao tamanho “b” e igual ao tamanho de “y”. Porém, observa-se que se existir um *pixel* no plano imagem (cartesiano) caindo sobre a borda (uma das quatro extremidades do plano imagem), a colocação de uma circunferência gerada com centro nela cairia parte fora do arranjo acumulador, como pode-se observar na figura 5.9.

FIGURA 5.9– Circunferência geradas com centro em *pixels* da imagem que ficam nas bordas.

Para evitar perda de informação, são definidas as dimensões do arranjo acumulador, como:

$$- raio \leq a \leq (x_{max} + raio) \text{ e } - raio \leq b \leq (y_{max} + raio) \quad (5.7)$$

ou

$$ARRANJO[X_{m\acute{a}x} + 2*raio] \times [Y_{m\acute{a}x} + 2*raio] \quad (5.8)$$

Tomando cada um dos cinco *pixels* da figura 5.7, obtém-se o traçado conforme a figura 5.8, que lançados no arranjo acumulador nos fornece valores entre 0 (zero) e 4 (quatro). O valor 4 é exatamente onde cruzam quatro das cinco circunferências traçadas, possuindo coordenadas P[(3,5),(3,5)].

Este método permite detectar circunferências no plano imagem, mesmo que esta seja descontínua ou incompleta, como mostra a figura 5.10:

FIGURA 5.10 - Circunferências fracionada e incompleta

Para se desenvolver um algoritmo que trace as circunferências no arranjo acumulador pode-se aplicar vários métodos. Um dos métodos consiste em utilizar a equação da circunferência diretamente; outro utiliza o algoritmo de “Bresenhan”. Em

qualquer dos métodos, o traçado da circunferência é realizado ou para um quadrante ou para um octante apenas e, em seguida, utiliza-se os mesmos valores de coordenadas para os demais quadrantes ou octetos por simetria.

KIMME *et al.* (1975) propõem a implementação de outro método utilizando Hough para detecção de circunferência da seguinte maneira: dado um conjunto de *pixels* em que se queira detectar uma circunferência, usando a indicação de DUDA & HART (1972) calcula-se para cada *pixel* da imagem digitalizada as operações $r * \cos \theta$, $r * \sin \theta$, para um dado r fixado, enquanto θ é incrementado de um valor $\Delta\theta$. Onde $\Delta\theta$ é função de r . Para todo o processamento, isto é, para todos os valores de raio adotados, o produto $r * \Delta\theta$ é sempre constante, permitindo que todo comprimento de arco entre dois *pixels* traçados no arranjo acumulador sejam iguais. Portanto, o processamento de cada circunferência exige um número de operações igual a $2\pi/\Delta\theta$. Como o tempo de processamento é significativo, para minimizá-lo Kimme (1975) sugeriu gradientes de direção para selecionar a porção de circunferência a ser traçada no arranjo acumulador. O processo consistiu em tomar um conjunto de *pixels* e avaliar a direção central deles, isto é, a direção do possível centro e traçar somente arcos de circunferência com seus limites voltados apenas para a posição central dos *pixels* da imagem. Conseqüentemente, ao invés de traçar uma circunferência total, traça-se os arcos com tamanhos pequenos, o que diminui o tempo de processamento e produz menos votos acumulados no arranjo acumulador. Após este procedimento, faz-se um “threshold” no arranjo acumulador e obtêm-se os possíveis centros da circunferência, com os quais encontram-se as circunferências da imagem.

5.2.1 Algoritmo da Transformada de Hough para objetos circulares

O algoritmo da Transformada de Hough para objetos circulares consiste basicamente de cinco passos (maiores detalhes são encontrados no Apêndice 2):

1- Primeiramente deve-se quantizar o espaço de parâmetros tendo como resultado um conjunto de valores que serão utilizados como parâmetro para os passos seguintes;

2- Zerar o arranjo acumulador $M(a,b)$;

3- Computar a magnitude do gradiente $G(x,y)$;

4- Para cada *pixel* de borda em $G(x,y)$, incrementar todos os *pixels* acumulados no arranjo $M(a,b)$ ao longo da linha seguindo:

$$a = x + r \cdot \cos \theta$$

$$b = y + r \cdot \sin \theta$$

5- Identificar o ponto máximo no arranjo acumulador, pois este corresponde ao centro do objeto circular presente na imagem.

A Figura 5.11 (b) mostra o resultado com a aplicação destas rotinas sobre a imagem em (a). Nota-se que são obtidas várias regiões candidatas a conterem os olhos, sendo que a próxima tarefa é tentar determinar quais destas regiões são os olhos.

FIGURA 5.11: - (a) Imagem de entrada; (b) Arranjo acumulador criado pela HT

5.2 Distância Euclidiana

Para encontrar o verdadeiro par de olhos, foi usada a Distância Euclidiana que é dada pela equação 5.9.

$$D_e(x,y) = \sqrt{(x_i - y_i)^2 + (x_j - y_j)^2} \quad (5.9)$$

FIGURA 5.12 – Proporção do cálculo da Distância Euclidiana entre os olhos e a largura da imagem.

A cada candidato a olho obtido do arranjo acumulador (picos de máximo do arranjo), é realizada uma varredura na imagem correspondente, 10 linhas acima e 10 linhas abaixo, para se localizar seu provável par. Quando o par existe, a distância entre eles é calculada (D_e). Para que seja possível definir que essa distância é a procurada entre os olhos, algumas considerações sobre biometria facial são realizadas, como por exemplo, proporção entre largura da imagem e distância medida, mostradas na Figura 5.12. Assim, os candidatos menos prováveis são eliminados confirmando-se apenas um par como a localização procurada. A D_e encontrada entre os olhos variou de 50 a 70 *pixels* sendo a largura da imagem de 200 *pixels*. Logo, a proporção foi de 1/3 ou 33%. Essa metodologia foi possível porque as imagens foram adquiridas de maneira padronizada, ou seja, mesma distância focal, iluminação, contraste e resolução espacial. A Figura 5.13 mostra uma imagem de face onde os dois olhos foram localizados e a proporção corretamente definida.

FIGURA 5.13 – Localização dos olhos pela proporção de 33%.

Capítulo 6

Resultados, Trabalhos Futuros e Conclusões

6.1 Resultados

Foram testadas 100 imagens, de um banco de imagens padrão para reconhecimento de faces, com resolução espacial de 200×200 *pixels*.

A Figura 6.1 mostra uma face com variação de expressão, podendo-se observar que os olhos são facilmente localizados, pois a íris mantém sua circularidade e a alteração da expressão não afeta o algoritmo.

Pequenas rotações no eixo que contém os olhos, ou no eixo perpendicular a este, também são toleradas como pode ser visto na Figura 6.2. Observa-se, entretanto, que a combinação dos métodos aqui utilizados (HT e Distância Euclidiana) obtém resultados superiores a outros métodos mesmo com a imagem possuindo os olhos fechados ou entreabertos. O algoritmo é robusto bastante para determinar sua localização, conforme mostra a Figura 6.3. A taxa de acerto obtida com todo o banco de imagens foi de aproximadamente 98%, considerando-se faces frontais, com bom contraste e baixo nível de ruído, condições estas facilmente conseguidas em sistemas automáticos de segurança ou identificação pessoal.

FIGURA 6.1 : Resultado da detecção de olhos com variação de expressão

FIGURA 6.2: Resultado da detecção de olhos com variação de rotação

FIGURA 6.3: Resultado da detecção de olhos com olhos entreabertos e fechados

O resultado do algoritmo proposto em nosso trabalho foi comparado com o método desenvolvido por LIN & FAN (2001). As três etapas desse método são descritas a seguir:

Etapas 1 - Pré-processamento da imagem de entrada para uma imagem binária

Nesta etapa, a imagem é convertida em uma imagem binária pela limiarização global simples com um limiar único T . Em seguida, como mostrado na Figura 6.4, é realizada a operação de abertura (erosão primeiro, depois dilatação) para remover ruídos, e então a operação de fechamento (dilatação primeiro, depois erosão) para eliminar buracos.

FIGURA 6.4 : (a) Imagem de entrada; (b) conversão da imagem para binária; (c) operação de abertura; (d) operação de fechamento

Etapa 2 - Rotular todos os componentes 4-conectados e encontrar o centro de cada bloco

Aqui é usado o “raster scanning” (da esquerda para a direita e de cima para baixo) para obter os componentes 4-conectados, rotulá-los e encontrar o centro de cada bloco, os quais são exibidos na Figura 6.5.

FIGURA 6.5 : Os pontos vermelhos representam os centros de cada bloco (ampliados 8x).

Etapa 3 - Encontrar quaisquer 3 centros de 3 diferentes blocos que formam um triângulo isósceles

Nesta etapa, procura-se por centros de blocos que formem um triângulo, baseado na regra “a combinação de dois olhos e uma boca” (Figura 6.6).

FIGURA 6.6 : Detecção dos olhos através da geometria do triângulo isósceles

Esse método é muito vulnerável a ruídos e sombra na imagem, como mostra a Figura 6.7. Na operação de abertura e fechamento, o olho uniu-se ao cabelo e ao fundo, formando um único bloco. Conseqüentemente, não haverá um centro neste olho, o que impedirá a localização correta.

FIGURA 6.7: Imagens com problemas na localização dos centros dos blocos devido a olhos conectados a cabelo e sobrancelhas.

Um outro problema com o método é o limiar utilizado. Com descrito na etapa 1, a imagem de entrada é convertida em uma imagem binária pela limiarização global simples com um limiar único T . Esse limiar não funciona corretamente para a maioria das imagens (Figura 6.8) e a realização de uma binarização imperfeita leva a consequências desastrosas na operação de abertura. Na figura 6.8 isto provocou o desaparecimento do nariz e da boca e, conseqüentemente, a localização do terceiro ponto triângulo isósceles.

FIGURA 6.8: (a) Imagem binarizada; (b) erosão e (c) dilatação para realização da operação de abertura, provocando eliminação da boca.

A Tabela 6.1 ilustra os resultados dos testes comparativos entre o método proposto neste trabalho - a HT e o método implementado por LIN & FAN (2001) - geometria do triângulo isósceles. Através dos resultados experimentais, constatou-se que a performance da HT é superior a do triângulo isósceles. Em um banco de imagens com baixo nível de ruído e bom contraste, a localização é realizada em 98% dos casos, mostrando boa confiabilidade, enquanto que o método do triângulo isósceles alcançou somente 36% de acerto. As imagens onde o triângulo isósceles falhou a HT funcionou são apresentadas no Apêndice 1, Seção 1.3 Constatou-se, também, que num Pentium III 750 MHz, o tempo de processamento requerido para o método de Hough foi menor que o requerido para o método do triângulo isósceles.

TABELA 6.1- Testes comparativos entre os métodos utilizados

	Índice de acerto	Tempo de processamento
Transformada de Hough	98%	7 s
triângulo isósceles	36%	10 s

6.3 Trabalhos Futuros e Conclusões

6.3.1 Conclusões

Os sistemas de reconhecimento automático e identificação pessoal são cada vez mais necessários e seu estudo é relativamente recente. Há três anos surgiram os primeiros sistemas comerciais para reconhecimento automático de faces. Não há ainda um algoritmo que se destaque, mas muitas pesquisas em processo de formação e consolidação de suas bases teóricas.

Para que se possa eficientemente reconhecer uma pessoa em uma imagem digital, é primordial a localização de características que disparem o processo de reconhecimento. Nesse sentido, os olhos são referências de grande importância, para que a partir deles possa se medir parâmetros biométricos essenciais.

A proposta deste trabalho de se utilizar a Transformada de Hough seguida da Distância Euclidiana baseada em parâmetros biométricos de faces humanas mostra-se de grande valia na localização dos olhos.

O Apêndice 1, Seção 1.1, mostra as imagens do banco de imagens utilizado e o resultado da localização dos olhos, marcadas através de uma cruz vermelha. Pode-se observar que o método funciona corretamente no banco considerado. Já a Seção 1.2 apresenta as imagens onde o algoritmo falhou. Na primeira imagem, o algoritmo não localizou os olhos devido a problemas na formação do arranjo acumulador, este ocasionado em virtude da presença de sombra e ruídos na aquisição. Na segunda imagem, o algoritmo localizou um arranjo acumulador na orelha direita, o qual, por coincidência, tem a mesma proporção dos olhos em relação à largura da imagem.

Em um banco de imagens com baixo nível de ruído e bom contraste, a localização é realizada em 98% dos casos, mostrando boa confiabilidade, rapidez e repetitividade. O algoritmo descrito faz parte de um programa de pesquisa mais amplo em desenvolvimento no Laboratório de Visão Computacional, visando a identificação automática em sistemas de segurança.

6.3.2 Trabalhos Futuros

Este trabalho de mestrado faz parte de um programa mais amplo em desenvolvimento, visando a identificação em sistemas de segurança e foi responsável pelo passo inicial para a solidificação do mesmo. A seguir são apresentadas algumas sugestões para outros trabalhos, que objetivam completar o sistema e aprimorar a implementação aqui desenvolvida:

- Desenvolvimento de um sistema de aquisição de imagens faciais, fixando determinados parâmetros importantes para a obtenção de imagens de melhor qualidade, tais como iluminação uniforme, distância fixa da câmera em relação à pessoa a ser fotografada;
- Implementação de outras técnicas para a localização do nariz e da boca. Extrair características e aplicá-las a classificadores para completar o sistema de identificação;
- Implementação da outra forma de utilização da transformada de Hough – i.e., algoritmo de “Bresenhan”- para tentar obter um ganho de velocidade no processo de localização dos candidatos a olhos.
- Possibilidade de aplicação dos dois métodos para tornar o programa mais robusto.

REFERÊNCIAS BIBLIOGRÁFICAS

- ANTONIE, J.; CARRETE, P.; MURENZI, R.; PIETTE, B. (1993). Image analysis with two-dimensional continuous wavelet transform. *Signal Processing*, vol. 31, pp. 241-272.
- BALLARD, D. H. (1981). Generalizing the Hough transform to detect arbitrary shapes. *Pattern Recognition*, vol. 13, nº 2, pp. 111-122.
- BANON, G. J. F. (1997). Morphological Approach for Template Matching. *Anais do X SIBGRAPI'97*, Artigo 22, pp. 171-178.
- BAUDEL, T; BEAUDOUIN-LAFON, M. C.(1993). Remote control of objects using free-hand gestures. *CAMC*, pp. 28-.
- BARTLETT, M. S.; HAGER, J. C.; EKMAN, P.; SEJNOWSKI, T. J. (1999). “Measuring facial expressions by computer image analysis”. *Psychophysiology*, vol. 36, pp. 253-264.
- BEN-TZVI, D.; NAQVI, A.; SANDLER, M. (1990). Synchronous multiprocessor implementation of the Hough transform. *Computer Vision, Graphics and Image Processing*, vol. 52, pp. 437-446.
- BEYMER, D. (1996). <http://www.odc.com.anthro/>, Antropometric Resource Center.
- BLACK, M. J.; YACOOB, Y. (1998). Recognising Facial Expressions in Image Sequences using Local Parameterised Models of Image Motion. *International Journal on Computer Vision*, vol. 25, nº 1, pp. 23-48.
- BRUNELLI, R.; POGGIO, T. (1993). Face Recognition: Features versus Templates. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, Vol. 15, nº 10, pp. 1042-1052.

- CAI, J.; GOSHTASBY, A. (1999). Detecting human faces in color images. *Image and Vision Computing*, Vol. 8, nº 1, pp. 63-75.
- CAMPOS, T.; FERIS, T.; CESAR, R. (2000). Eigenfaces versus eigeneyes: first steps towards performance assessment of representations for face recognition. *Automatic Systems for the Identification and Inspection of Humans*, vol. 2277, SPIE.
- CASASENT, D.; KRISHNAPURAM, R. (1987). Curved object location by Hough transformation and inversions, *Pattern Recognition*, vol. 20, nº 1, pp. 181-188.
- CHAN, R.; SIU, W. C. (1991). New Parallel Hough transform for circles, *IEEE Proceedings-E*, vol. 138, nº 05.
- CHELLAPPA, R.; WILSON, C. L.; SIROHEY., S. (1995). Human and Machine Recognition of Faces: A Survey. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, Vol. 83, nº 5.
- CHEN, Q.; WU, H.; YACHIDA, M. (1995). Face detection by fuzzy pattern matching. *Proceedings of the Fifth International Conference on Computer Vision*, pp. 591-596.
- CIPOLLA, R.; OKAMOTO, Y.; KUNO, Y. (1993). Robust structure from motion using parallax. *IEEE Proceedings on International Conference on Computer Vision*.
- CHOW, G.; LI, X. (1993). Toward a system for automatic facial feature detection, *Pattern Recognition*, vol. 29, nº 5, pp. 1739-1755.

- COOTES, T. F.; TAYLOR, C. J. (1996). Locating faces using statical feature detectors. *Proceedings of the Second Internacional Conference on Automatic Face and Gesture Recognition*, pp. 204-209.
- CROWLEY, J. L.; BERARD, F. (1997). Multi-modal tracking of faces for video communications. *Proceedings of the Computer Vision and Pattern Recognition*, Porto Rico, n° 6, pp. 640-645.
- CYGANSK, D. (1996). <http://xfactor.wpi.edu/Works/MQP/avi96/mqp/node42.html>. Template Matching.
- DAI, Y.; NAKANO, Y. (1996). Face-texture model based on SGLD and its applications in face detection in a color scene. *Pattern Recognition*, Vol. 29, n° 6, pp. 1007-1071.
- DAI, Y.; NAKANO, Y. (1998). Recognition of facial images with low resolution using a hopfield memory model. *Pattern Recognition*, Vol. 31, n° 2, pp.159-167, 1998.
- DARELLE, T.; PENTLAND, A. (1993). Space-times gestures. *IEEE CVPR*, pp. 335-340.
- DAVIES, E. R. (1992). Modelling peak shapes obtained by Hough transform, vol. 130, n° 1, 1992.
- DUCOTTET, C.; DANIERE, J.; MOINE, M.; SCHON, J. P.; COURBON, M. (1994). Localization of objects with circular symmetry in a noisy image using Wavelet Transforms and adapted correlation. *Pattern Recognition*, vol. 27, n° 3, pp. 351-364.
- DUDA, R.O.; HART, P.E. (1972). Use of the Hough transformation to detect lines and curves in pictures. *Graphics and Image Processing*, vol.12, n° 1, pp.11-15.

- EKMAN, P.; FRIESEN, W. V. (1978). Facial Action Coding System (FACS): Manual. *Consulting Psychologists Press*, Palo Alto.
- ESSA, I. A; PENTLAND, A. P. (1997). Coding analysis interpretation and recognition of facial expressions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.19, n° 7, pp. 757-763.
- FERIS, R.; CAMPOS, T.; CESAR, R. (2000). Detection and tracking of facial features in video sequences. In *Lecture Notes in Artificial Intelligence*, vol. 1793, pp. 127-135, Acapulco, Mexico.
- FUKUI, K.; KAWAGUCHI, O. (1997). Facial feature point extraction method based on combination of shape extraction and pattern matching, *IEICE Trans.*, vol. J80-DII, pp. 2170-2177.
- FUKUMOTO, M.; MASE, K.; SUENAGA, Y. (1992). Real-time detection of pointing actions for a glove-free interface. *IAPR Workshop on Machine Vision Applications*, pp. 473-476.
- GARCIA, C.; ZIKOS, G.; TZIRITAS, G. (2000). Wavelet packet analysis for face recognition. *Image and Vision Computing*, vol. 18, n° 4, pp. 289-297.
(NUMERO)
- GERIG, G.; KLEIN, F. (1986). Fast contour identification through efficient Hough transform and simplified interpretation strategy, *IEEE*.
- GONZALEZ, R. C., WOODS, R. E. (1993). Digital Image Processing. *Addison-Wesley Publishing Company, Inc.*
- GOVINDARAJU, V. (1996). Locating human faces in photographs. *Internacional Journal of Computer Vision*, vol 19, n° 2, pp. 129-146.

GROß, M. (1994). *Visual Computing*, Springer-Verlag.

HALLINAM, P. W. (1991). "Recognizing human eyes", *SPIE Proceedings*, vol. 1570, Geometric Methods in Computer Vision, 1991, pp. 214-226.

HAN, C. C; LIAO, H. Y. M. ; YU, G. J. ; CHEN, L. H. (2000). Fast face detection via morphology-based pre-processing, *Pattern Recognition*, vol. 33, n° 10, pp. 1701-1712.

HAN, J. H.; LÁSZLÓ, T.; POSTON, T. (1994). Fuzzy Hough Transform. *Pattern Recognition Letters*, vol. 15, pp. 649-658.

HANAHARA, K.; HIYANE, M. (1990). A circle-detection algorithm simulating wave propagation, *Machine Vision and Applications*, vol. 3, pp.97-111.

HUANG, J.; LIU, C.; WECHSLER, H. (1998). Eye Detection and Face Recognition Using Evolutionary Computation, *Proceeding of NATO-ASI on Face Recognition*, pp.348-377.

HUANG, T; PAVLOVIC, V. (1995). Hand gesture modeling, analysis, and synthesis. *Proceedings of International Workshop on Automatic Face and Gesture Recognition*, pp. 73-79.

HUECKEL, M. (1973). A local visual operator which recognizes edges and lines. *J. Ass. Comp.*, vol. 20, pp. 634-646.

HUSSAIN, Z. (1991). Digital Image Processing, Pratical Applications of Parallel Processing Techniques. *Ellis Horwood Limited*, West Sussex.

ILLINGWORTH, J.; KITTLER, J. (1988). A survey of the Hough transform, *Computer Vision, Graphics and Image Processing*, vol. 44, pp. 87-116.

- ILLINGWORTH, J.; KITTLER, J. (1987). The adaptive Hough transform. *IEEE Trans. Pattern Analysis Machine Intelligence*, vol. 9, nº 5, pp. 690-695.
- KANADE, T.; COHN, J. F.; TIAN, Y. (2000a). Comprehensive database for facial expression analysis". *IEEE Proceedings of Fourth International Conference on Automatic Face and Gesture Recognition*, pp. 46 –53.
- KANADE, T.; COHN, J. F.; TIAN, Y. (2000b). Recognizing actions units for facial expressions analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, nº 2, fevereiro, pp. 97-115.
- KAWAGUCHI, T; HUDAKA, D; RIZON, M.(2000). Robust extraction of eyes from faces, *Proc. 15th International Conference on Pattern Recognition*, vol. 1, pp. 1109-1114.
- KIMME, C.; BALLARD, D.; SKLANSKY, J. (1975). Finding circles by an array of accumulators. *Communications of the ACM*, vol. 18, nº2, pp.120-122.
- KIRBY, M. (1990). Application of the Karhunen-Lòeve Procedure for the Chacterization of Human Faces. *IEEE Transactions on Pattern Analysis an Machine Intelligence*, vol. 12, nº 1, pp. 103-108.
- KIRBY, M.; SIROVICH, L. (1990). Application of the karhunen-Loève procedure for the characterization of human faces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, nº 1, pp. 103-108.
- KOBAYASHI, H.; HARA, F. (1997). Facial interaction between animated 3D face robot and human beings. *IEEE International Conference on System, Man and Cybernetics*, pp. 3732-3737.
- KOSKO, B. (1992). Neural Network and Fuzzy Systems. *Prentice Hall*.

- KRUGER, V.; SOMMER, G. (2000). Gabor wavelet networks for object representation. *Technical Report 2002*, University of Kiel, Germany.
- LAM, K.; LI, Y. (1998). An efficient approach for facial feature detection, *Proceedings on Fourth International Conference Signal Processing*, vol. 2, pp. 1100–1103.
- LAM, K. M.; YAN, H. (1998). An analytic-to-holistic approach for face recognition based on a single frontal view, *IEEE Trans. on PAMI*, vol. 20, n° 7, pp. 6673-686.
- LEE, E.; WHALEN, T. (1993). Computer Image Retrieval by Features: Suspect Identification. *Proceedings of Interchi'93*, n° 4, pp.494-499.
- LI, H.; LAVIN, M. A.; MASTER, R. J. L. (1986). Fast Hough Transform: A hierarchical approach. *Computer Visio*, Graphics and Image Processing, vol. 36, pp. 139-161.
- LIN, C.; FAN, K. C. (2001). Triangle based approach to the detection of human face, *Pattern Recognition*, Vol. 34, pp. 1271-1284.
- LOW, A.(1991). Introductory computer vision and image processing. *McGraw-hill book company*.
- MERLIN, P. M.; FARBER, D. J. (1975). A parallel mechanism for detecting curves in pictures. *IEEE Trans. Computer*, vol. 24, pp. 96-98.
- MIYAKE, Y.; SAITOH, H.; YAGUCHI, H., TSUKADA. N. (1996). Facial pattern detection and color correction from television picture and newspaper printing. *Journal of Imaging Technology*, vol 16, n° 5, pp.165-169.

- MOGHADDAM, B. (1995). An Automatic System for Model-Based Coding of Faces. *M.I.T. Media Laboratory Perceptual Computing Section Technical Report*, nº 317.
- MORIMOTO, C. H.; YACOOB, Y.; DAVIS, L. (1996). Recognition of head gestures using hidden markov models. *ICPR*, Vienna, Áustria.
- NASA (1996). Best Template Matching. <http://www.msfc.nasa.gov/irgrp.bestmatch.html>, Nasa/Marshall Space Flight Center.
- NIXON, M. (1985). Eye spacing measuring for facial recogniton. *SPIE Proc.*, Vol. 575, pp.279-285.
- O'GORMAN, F. O.; CLOWES, M. B. (1976). Finding picture edges through collinearity of feature points. *IEEE Trans. Computer*, vol. 25, pp. 449-456.
- PANTIC, M.; ROTHKRANTZ, L. J. M. (2000). Expert systems for automatic analysis of facial expressions. *Image and Vision Computing*, vol. 18, pp. 881-905.
- PARKER, J. R. Practical Computer Vision Using C, *John Wiley & Sons, Inc.*, New York,U.S.A., 1994.
- PENTLAND, Alex (2000). Looking at people: Sensing for ubiquitous and wearable computing. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, nº 1, pp. 107-119.
- PERELMUTER, G; CARRERA, E. (1995). Reconhecimento de Imagens Bidimensionais Utilizando Redes Neurais Artificiais. *Anais do VIII Sibgrapi*, pp. 197-203.

- ROSENFELD, A. (1988). Hough transform algorithm for mesh-connected SIMD parallel processor. *Computer Vision, Graphics and Image Processing*, vol. 41, pp. 293-305.
- ROMDHAMI, S. (1997). <http://www.elec.gla.ac.uk/~romdhami>. Face Recognition Using Principal Component Analysis.
- ROMDHAMI, S. (1996). *Face recognition using principal component analysis*. Dissertação de Mestrado, Departamento de Engenharia Elétrica e Eletrônica, Universidade de Glasgow, UK.
- ROWLEY, H. A.; BALUJA, S.; KANADE, T. (1998). Neural Network-based Face Detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, nº 1, pp. 23-28.
- SABER, E.; TEKALP, A. M. (1998). Front view-face detection and facial feature extraction using color, shape and symmetry based cost functions. *Pattern Recognition Letters*, Vol. 19, nº 8, pp. 669-680.
- SAHOO, P. K.; SOLTANI, S.; WONG, A. K. C (1988). An Survey of Thresholding Techniques. *Computer Vision, Graphics and Image Processing*, nº 41, pp. 233-260.
- SAMAL, A.; IYENGAR, P. A. (1992). Automatic Recognition and Analysis of Human Faces and Facial Expressions: A Survey. *Pattern Recognition*, vol. 25, pp. 65-77.
- SANTINI, S.; JAIN. R. (1999). Similarity Measures. *IEEE Transactions on Pattern Recognition and Machine Intelligence*, Vol 21, nº 9.
- SHAPIRO, S. D. (1975). Transformation for the computer detection of curves in noisy pictures. *Comput. Graphics Image Process*, vol. 4, pp. 328-338.

- SHAPIRO, S. D. (1978a). Feature space transforms for curve detection. *Pattern Recognition*, vol. 10, pp. 129-143.
- SHAPIRO, S. D. (1978b). Properties of transforms for the detection of curves in noisy pictures, *Comput. Graphics Image Process*, vol. 8, pp. 219-236.
- SCHIELE, B.; WAIBEL. (1995). Gaze tracking based on face-color. *Internacional Workshop on Face and Gesture Recognition*, Zurich.
- SILVA, L.C. De; NG, P. C. (2000). Bimodal emotion recognition. *IEEE Proceedings of Fourth International Conference on Automatic Face and Gesture Recognition*, pp. 332-335.
- SIROHEY, S. A., ROSENFELD, A. (2001). Eye detection in a face image using linear and nonlinear filters, *Pattern Recognition*, vol. 34, n° 7, pp. 1367-1391.
- SMERALDI, F.; CARMONA, O.; BIGÜN, J. (2000). Saccadic search with Gabor features applied to eye detection and real-time head tracking. *Image and Vision Computing*, Vol. 18, n° 4, pp.323-329.
- SOBOTTA, K.; PITAS, I. (1996). Segmentation and tracking of faces in color images. *Proceedings of the Second Internacional Conference on Automatic Face and Gesture Recognition*, pp. 236-241.
- SUNG, K.K; POGGIO, T. (1998). Example-based learning for view-based human face detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, n° 1, pp. 39-51.
- THALMANN, N.M.; KALRA, P.; ESCHER, M. (1998). Face to virtual face. *Proceedings of the IEEE*, vol 86, n° 5, pp. 870-883.

- TOLBA, A.S.; ASHRAF, M.; ABU-REZQ. (2000). Eyes-strip extraction for detection of a human face, ICM '99. *The Eleventh International Conference on Microelectronics*, pp.109 –112.
- TURK, A. M.; PENTLAND, A. (1991). Face recognition using eigenfaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 15, n° 10, pp.1042-1052.
- UCHIYAMA, T.; ARBIB, M. A. (1994). Color image segmentation using competitive learning. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 16, n° 12, pp. 1197-1206.
- VALENTIN, D. (1994). Connectionist Models of Face Processing: A Survey. *Pattern Recognition*, vol. 27, n° 1, pp. 1209-1230.
- WANG, J.; TAN, T. (2000). A new face detection method based on shape information. *Pattern Recognition Letters*, vol. 21, pp. 463-471.
- WEI, G.; SETHI, K. (1999). Face detection for image annotation. *Pattern Recognition Letters*, vol. 20, n° ?, pp. 1313-1321.
- YANG, G., HUANG, T. (1994). Human face detection in a complex background. *Pattern Recognition*, vol. 27, n° 1, pp. 53-63.
- YANG, J.; LU, W.; WAIBEL, A. (1997). Skin-color modeling and adaptation. *Technical Report*, CMU-CS-97-146, maio.
- YOON, K. S.; HAM, Y. K.; PARK, R. (1998). Hybrid approaches to frontal view face recognition using the hidden markov model and neural network. *Pattern Recognition*, Vol. 31, n° 3, pp. 283-293.

- YESCHURUN, Y.; REISFELD, D.; WOLFSON, H. (1991). Symmetry: a context free cue for foveated vision, *Neural Networks for Pattern Recognition*, Vol. I, H. Wechsler, ed., Academic Press.
- YUEN, H. K.; PRINCEN, J.; ILLINGWORTH, J.; KITTLER, J. (1990). Comparative study of Hough transform methods for circle finding, vol. 8, n° 1, pp. 71-77.
- YUILLE, A; COHEN, D.; HALLINAN, P. (1992). Feature extraction from faces using deformable templates. *Int. J. Computer Vision*, vol. 8, n° 2, pp. 99-111.
- YUILLE, A; COHEN, D.; HALLINAN, P. (1989). Feature extraction from faces using deformable templates, *IEEE Computer Soc. Conf. on Computer Vision and Pattern Recognition*, pp. 104-109.
- YOW, K. C.; CIPOLLA, R. (1997). Feature-based human face detection. *Image and Vision Computing*, Vol. 15, pp. 713-735.
- ZANGH, Q.; BENVISTE, A. (1992). Wavelets networks. *IEEE Transactions on Neural Networks*, vol. 3, n° 6, pp. 889-898, novembro.
- ZLOCHOWER, A. J. (1997). Deciphering emotion from the face: An evaluation of FACS, EMG, and computer-vision based approaches to facial expression analysis. *Universidade de Pittsburgh*.

APÊNDICE 1 – Imagens do banco de imagens usado para testes

Seção 1.1 - Localização correta dos olhos

Seção 1.2 - Falha na localização dos olhos pela HT

Seção 1.3 - Imagens onde o método do triângulo isósceles falhou e a HT funcionou

APÊNDICE 2 - Listagem das rotinas implementadas

Seção 2.1 Implementação do algoritmo da HT para objetos circulares

Passo_1

```
FOR teta:=0 TO 359 DO
  BEGIN
    cosseno[teta] := cos(teta/180*pi);
    seno[teta] := sin(teta/180*pi);
  END;
```

Passo_2

```
FOR x:=0 TO 399 DO // imagem 200 x 200 com bordas iguais a 100 (399)
  FOR y:=0 TO 399 DO
    FOR z:=5 TO 7 DO
      acumuc[x,y,z] :=0; // inicializa o acumulador com zeros
```

Passo_3

```
FOR r:=5 TO 7 DO // calcula HT com raios entre 5 e 7
  BEGIN
    dteta := trunc(30/r);
    FOR x:=1 TO imagebig2.height-2 DO
      FOR y:=1 TO imagebig2.width-2 DO
        BEGIN
          cor := imagebig.Canvas.Pixels[x,y];
          Rcor := getrvalue(cor);
```

Passo_4

```
IF Rcor<>0 THEN
  BEGIN
    teta := 0;
    WHILE teta < 360 DO
```

```

BEGIN
    a := trunc(x + r * coseno[teta]);
    b := trunc(y + r * seno[teta]);
    acumuc[100+a,100+b,r]:=acumuc[100+a,100+b,r]+1;
    teta := teta + dteta;
END;
END;

```

Passo_5

```

FOR y:=0 TO 399 DO
    FOR x:=0 TO 399 DO
        BEGIN
            IF acumuc[x,y,r] > 35 THEN
                BEGIN
                    vy[ind].x:=x-100;
                    vy[ind].y:=y-100;
                    vy[ind].r:=acumuc[x,y,r];
                    inc(ind);
                END;
            END;
        END;
    END;

```

Seção 2.2 - Localização dos olhos pelo cálculo proporção da Distância Euclidiana entre os olhos e a largura da imagem

```

FOR y:=0 TO ind-1 DO // imagem 200 x 200 com bordas iguais a 100 (399)
    FOR x:=y+1 TO ind DO // raios maximos = 100 (399)
        BEGIN
            DE := sqrt(sqrt(vy[y].x - vy[y].y)+ sqrt(vy[x].x - vy[x].y));
            IF ((largura / 4)>=DE) AND (largura/4 * 1.33<=DE) AND (not achou)
            AND ((vy[y].y - vy[x].y)< 10) AND ((vy[y].y - vy[x].y)> -10) THEN
                BEGIN
                    vyx:= vy[y].x; //coluna
                    vyy:= vy[y].y; //linha
                    vxy:= vy[x].y; //coluna
                    vxx:= vy[x].x; //linha
                    DE1:=DE;
                    d:= 1;
                END;
            END; // fim da coluna
        END;
    END;

```

```
IF D=1 THEN //condição para não desenhar duas vezes as cruzinhas
BEGIN
  imagebig3.Canvas.Pen.Color := clRed;
  imagebig3.Canvas.MoveTo(vyx-10,vyy);
  imagebig3.Canvas.LineTo(vyx+10,vyy);
  imagebig3.Canvas.MoveTo(vyx,vyy-10);
  imagebig3.Canvas.LineTo(vyx,vyy+10);
  imagebig3.Canvas.MoveTo(vxx-10,vxy);
  imagebig3.Canvas.LineTo(vxx+10,vxy);
  imagebig3.Canvas.MoveTo(vxx,vxy-10);
  imagebig3.Canvas.LineTo(vxx,vxy+10);
  d:= 2;
END;
```

Implementação do método do triângulo isósceles

```

PROCEDURE TForm1.TIssceles1Click(Sender: TObject);
VAR
  Dij, Djk, fregra,fregraT, Dik, Dik2, Dij2,max, reg1, reg2: real;
  xi, xj, xk, yi, yj, yk, ii, ij, ik, i,j,k : integer;

BEGIN
  reg1:=0;
  reg2:=0;
  fregra:=0;
  FOR i:=1 to nb do // nb = número de blocos encontrados
    FOR j:=1 to nb do
      FOR k:= 1 to nb do
        BEGIN
          If (i<>j) AND (i<>k) AND (j<>k) THEN
            BEGIN
              xi:= xc[i];
              yi:= yc[i];
              xj:= xc[j];
              yj:= yc[j];
              xk:= xc[k];
              yk:= yc[k];
              Dij:= sqrt(sqr(xi - xj) + sqr(yi-yj));
              Djk:= sqrt(sqr(xj - xk) + sqr(yj-yk));
              Dik:= sqrt(sqr(xi - xk) + sqr(yi-yk));
              max:= Dij;
              IF Djk > max THEN
                max:= Djk;

              reg1:= 0.25 * max - abs(Dij - Djk);
              reg2:= 0.25 * max - abs(Dij - Dik);
              fregraT:= reg1 + reg2; //fregra = (reg1 + reg2)/abs(reg1-reg2);
              IF (fregraT>fregra) AND ((Dik>60) AND (Dik<80)) AND ((yi<yj)
                AND (yk<yj)) AND (abs(Dij - Djk) <10) then
                BEGIN
                  fregra:=fregraT; //regra atual
                  ii:= i;      // ii, ij, ik: índices do vetor de centros
                  ij:= j;
                  ik:= k;
                  Dik2:= Dik;
                  Dij2:=(Dij+Djk)/2;
                END;
            END;
          END;
        END;
      END;
    END;
  END;

```

```
imagebig.canvas.pen.color:= clred;  
imagebig.canvas.moveto(xc[ii],yc[ii]);  
imagebig.canvas.lineto(xc[ik],yc[ik]);  
imagebig.canvas.lineto(xc[ij],yc[ij]);  
imagebig.canvas.lineto(xc[ii],yc[ii]);  
imagebig2.canvas.pen.color:= clred;  
imagebig2.canvas.moveto(xc[ii],yc[ii]);  
imagebig2.canvas.lineto(xc[ik],yc[ik]);  
imagebig2.canvas.lineto(xc[ij],yc[ij]);  
imagebig2.canvas.lineto(xc[ii],yc[ii]);  
END;
```