Processamento e análise de imagens para medição de vícios de refração ocular

MSc. Antonio Valerio Netto

Orientadora: *Profa. Dra. Maria Cristina Ferreira de Oliveira*

Co-orientador: *Prof. Dr. André Ponce de Leon F. de Carvalho*

Tese apresentada ao Instituto de Ciências Matemáticas e de Computação - ICMC-USP, como parte dos requisitos para obtenção do título de Doutor em Ciências – Ciências de Computação e Matemática Computacional.

“VERSÃO REVISADA APÓS A DEFESA”

Data da Defesa:	18/08/2003
Visto da Orientadora:	

USP – São Carlos
Setembro/2003

“O rio atinge seu objetivo porque aprendeu a desviar dos obstáculos”

“Quem não sabe o que procura, não pode interpretar o que encontra”

“É muito melhor arriscar coisas grandiosas, alcançar triunfos e glória, mesmo expondo-se a derrota, do que formar fila com os pobres de espírito que nem gozam muito nem sofrem muito, porque vivem nessa penumbra cinzenta, que não conhece vitória nem derrota”

T. Roosevelt (ex-presidente dos EUA)

Dedicatória

Á minha amada esposa Danielle pelo apoio e carinho recebido ao longo de todos estes anos de convivência. Aos meus familiares, amigos e, principalmente, ao senhor bom Deus que me ilumina todo tempo e em todo lugar.

Agradecimentos

A minha orientadora, Profa. Dra. Maria Cristina Ferreira de Oliveira, pelo bom convívio, apoio completo ao projeto e também por acreditar que poderíamos realizá-lo começando da estaca zero.

Ao Prof. Dr. André Ponce de Leon F. de Carvalho, meu co-orientador, pela amizade e pelas suas importantes contribuições na área de redes neurais artificiais, máquinas de vetores suporte e combinação de classificadores.

Ao pessoal de iniciação científica, Osvaldo César de Almeida, Rodrigo Oshiro e Giampaolo Libralão, meu muito obrigado por contribuírem diretamente no projeto e pela amizade que foi gerada ao longo do trabalho.

Ao Prof. Dr. João do Espírito Santo B. Neto pela amizade, pela força dada ao projeto e pelas suas contribuições na área de processamento de imagens e ao Prof. Dr. Alexandre C. B. Delbem pela amizade e auxílio na parte de extração de características e inteligência computacional.

À Profa. Dra. Agma J. Traina pelo apoio ao projeto, principalmente no período em que foi minha orientadora substituta. Muito obrigado.

Ao Prof. Dr. Evandro E. S. Ruiz da FFCLRP/USP pelas dicas sobre processamento de imagens médicas, ao Prof. Dr. Sérgio Muniz do Departamento de Física da UFSCar pela amizade e auxílio na área de física óptica, e ao Doutorando Estéfane Lacerda do Departamento de Informática da UFPE pela ajuda com as redes RBF com AG.

Agradecimento especial ao Prof. Dr. Larry Thibos e sua equipe da Escola de Optometria da Universidade de Indiana (EUA), pelas informações da área de Optometria, pela liberação do banco de imagens de Hartmann-Shack, e pelo bom convívio no período em que estive lá. *Nikole Himebaugh, Kevin Haggerty, Xin Hong, Mike Venable, Junle Qu and Prof. Arthur Bradley, thanks for everything. Dear Prof. Thibos thanks to believe in my project.*

Aos amigos e colegas do Grupo de Computação Gráfica e Processamento de Imagens do Laboratório de Computação de Alto Desempenho (LCAD), André, André

Toscano, Xavier, Luciano, Rodrigo, Viviane, Kleber, Alex, Lúcio, Helton, Billy, Igor, Veridiana, e Prof. Dr. Onofre Trindade. Ao pessoal do Laboratório de Inteligência Computacional (LABIC), Lorena, Valéria, Patrícia, Claudinha, Jaqueline, Gedson e Katti; e ao pessoal do Grupo de Bases de Dados e Imagens (GBdI), Camilo e César.

A todos os funcionários e outros professores do Departamento de Ciência de Computação e Estatística (SCE) que diretamente ou indiretamente contribuíram para o resultado final desse trabalho, em especial a Beth e a Tatiana. Muito obrigado.

Quero também agradecer a um grupo que teve papel fundamental no desenvolvimento desse projeto; tenham certeza que seus nomes não sairão da minha mente tão cedo... Obrigado fapespoftal, furious, mardita, tranquilo, transglobe, slash, k2, smac, gensmac, reynolds, oldroyd, maxwell, freeflow, spp3, dupla, goiaba e kilimanjaro.

À FAPESP pelo suporte financeiro.

Índice

RESUMO	1
ABSTRACT	2
1. INTRODUÇÃO	3
1.1 OBJETIVO E MOTIVAÇÕES	4
1.2 VISÃO GERAL DO SISTEMA.....	6
1.3 ORGANIZAÇÃO DA TESE	8
2. O OLHO HUMANO E OS VÍCIOS REFRACTIVOS.....	10
2.1 ESTRUTURA E FUNCIONAMENTO DO GLOBO OCULAR.....	10
2.2 ERROS DE REFRAÇÃO	12
2.3 PROCESSOS DE MEDIÇÃO DOS VÍCIOS REFRACTIVOS	14
2.3.1 RETINOSCOPIA.....	15
2.3.2 OPTÔMETROS.....	16
2.3.3 REFRACTORES AUTOMÁTICOS.....	19
3. AQUISIÇÃO DE DADOS DO OLHO HUMANO	21
3.1 A TÉCNICA DE HARTMANN-SHACK NA ÁREA OFTALMOLÓGICA	21
3.2 PRINCÍPIO DE FUNCIONAMENTO	24
3.3 SISTEMA DE AQUISIÇÃO.....	29
3.4 ANÁLISE DA IMAGENS BASEADA NO POLINÔMIO DE ZERNIKE.....	32
4. TÉCNICAS DE EXTRAÇÃO DE CARACTERÍSTICAS.....	37
4.1 ANÁLISE DE COMPONENTES PRINCIPAIS (PCA).....	38
4.2 TRANSFORMADA WAVELET DE GABOR.....	42
5. TÉCNICAS DE APRENDIZADO DE MÁQUINA PARA ANÁLISE DE IMAGENS. 46	
5.1 REDES NEURAIS ARTIFICIAIS	47
5.1.1 FUNDAMENTOS.....	48

5.1.2 <i>MULTILAYER PERCEPTRON (MLP) E RADIAL BASIS FUNCTION (RBF)</i>	51
5.2 RNAs PARA RECONHECIMENTO E ANÁLISE DE IMAGENS	54
5.3 ALGORITMOS GENÉTICOS ASSOCIADOS ÀS RNAs	59
5.4 MÁQUINAS DE VETORES SUPORTE.....	60
6. PROCESSAMENTO E EXTRAÇÃO DE CARACTERÍSTICAS.....	64
6.1 BANCO DE IMAGENS	64
6.2 PRÉ-PROCESSAMENTO.....	66
6.2.1 <i>CENTRALIZAÇÃO E PADRONIZAÇÃO DO CÍRCULO DE HARTMANN-SHACK</i>	67
6.2.2 <i>ELIMINAÇÃO DE RUÍDOS</i>	71
6.3 EXTRAÇÃO DE CARACTERÍSTICAS	73
6.3.1 <i>VETORES DE CARACTERÍSTICAS UTILIZANDO PCA</i>	73
6.3.2 <i>VETORES DE CARACTERÍSTICAS UTILIZANDO GABOR</i>	75
6.3.3 <i>COMPARAÇÃO ENTRE OS VETORES DE CARACTERÍSTICAS</i>	76
7. RECONHECIMENTO E ANÁLISE.....	78
7.1 CONSIDERAÇÕES SOBRE O BANCO DE IMAGENS	79
7.2 EXPERIMENTOS UTILIZANDO PCA E SVM.....	83
7.2.1 <i>VARIANDO A QUANTIDADE DE ATRIBUTOS DE ENTRADA</i>	87
7.2.2 <i>BALANCEAMENTO ARTIFICIAL DO BANCO DE IMAGENS</i>	88
7.3 EXPERIMENTOS UTILIZANDO GABOR E SVM.....	89
7.4 EXPERIMENTOS UTILIZANDO GABOR E REDES RBF	92
8. CONCLUSÕES	97
8.1 CONTRIBUIÇÕES E DISCUSSÃO.....	97
8.2 SUGESTÕES PARA TRABALHOS FUTUROS.....	100
REFERÊNCIAS BIBLIOGRÁFICAS.....	103
BIBLIOGRAFIA CONSULTADA	114
ANEXO 1 – COMPARAÇÃO DOS VETORES DE CARACTERÍSTICAS.....	119
ANEXO 2 – GRÁFICOS DOS RESULTADOS DOS EXPERIMENTOS COM GABOR E SVM	123

Lista de Figuras

FIGURA 1.1- VISÃO GERAL DO SISTEMA DE MEDIÇÃO DOS VÍCIOS REFRACTIVOS.	6
FIGURA 2.1- OLHO HUMANO E SEUS PRINCIPAIS COMPONENTES [ANA00].	11
FIGURA 2.2 - COMO UMA LENTE PODE SER UTILIZADA PARA CORRIGIR O ERRO REFRACTIVO CAUSADO PELA MIOPIA [DOE99].	13
FIGURA 2.3 - COMO UMA LENTE PODE SER UTILIZADA PARA CORRIGIR O ERRO REFRACTIVO CAUSADO PELA HIPERMETROPIA [DOE99].	14
FIGURA 2.4 - LUZ PROVENIENTE DA RETINA DO PACIENTE PASSANDO PELO ORIFÍCIO CENTRAL DO ESPELHO E ATINGINDO A RETINA DO EXAMINADOR [VEN95].	15
FIGURA 2.5 - ESQUEMATIZAÇÃO DO MODO COMO O RETINOSCOPISTA VERÁ O “REFLEXO” DA LUZ NA RETINA SE SOBREPONDO À FAIXA DE ILUMINAÇÃO, EM TERMOS DA DINÂMICA DO MOVIMENTO [VEN95].	16
FIGURA 2.6 - ESQUEMATIZAÇÃO DO PRINCÍPIO DE SCHEINER [VEN95].	17
FIGURA 2.7 - VISTAS DE UM AUTO-REFRATOR COMERCIAL [CAN97] [CAN99].	20
FIGURA 3.1 - PRINCÍPIO DO DISCO DE SCHEINER [THI00A].	25
FIGURA 3.2 - OPTÔMETRO QUE UTILIZA O PRINCÍPIO DE SCHEINER [THI00A].	25
FIGURA 3.3 - VISÃO GERAL DO <i>ABERROMETER</i> DE HARTMANN [THI00A].	26
FIGURA 3.4 - DESCRIÇÃO DE <i>FUNDUS CAMERA</i> CONVENCIONAL [THI00A].	27
FIGURA 3.5 - DESCRIÇÃO DE UM <i>FUNDUS CAMERA</i> MODIFICADO [THI00A].	27
FIGURA 3.6 - PRINCÍPIO DA ANÁLISE DA FRENTE DE ONDA EM UM OLHO EMÉTROPE [THI99A].	28
FIGURA 3.7 - PRINCÍPIO DA ANÁLISE DA FRENTE DE ONDA EM UM OLHO AMÉTROPE [THI99A].	28
FIGURA 3.8 - CONFIGURAÇÃO FINAL DO “ <i>ABERROMETER</i> ” HARTMANN-SHACK [THI00A].	29
FIGURA 3.9 - DIAGRAMA ESQUEMÁTICO DA CONFIGURAÇÃO EXPERIMENTAL UTILIZADO POR LIANG [LIA94].	30
FIGURA 3.10 - IMAGEM PADRÃO CORRESPONDENTE A ONDA PLANA MARCADA COM A ÁREA ÚTIL [LIA94].	31
FIGURA 3.11 - IMAGENS DO OLHO DE UM PACIENTE L (L1 E L2, RESPECTIVAMENTE) OBTIDAS EM DOIS DIAS CONSECUTIVOS [LIA94].	32
FIGURA 3.12 - IMAGENS DO OLHO DE UM PACIENTE B (B1 E B2, RESPECTIVAMENTE), OBTIDAS NO MESMO DIA COM DIFERENÇA DE MEIA HORA ENTRE UM E OUTRO [LIA94].	32
FIGURA 4.1 - COMPONENTE REAL DA FUNÇÃO DE GABOR NO DOMÍNIO, ESPACIAL (A) E DAS FREQUÊNCIAS (B) [DAU95].	44

FIGURA 5.1 - TAREFAS DE UMA CADEIA DE PROCESSAMENTO DE IMAGENS [EGM02]	55
FIGURA 6.1 - ARQUITETURA DO SISTEMA DE AQUISIÇÃO BASEADO NA TÉCNICA DE HS [VIS01].	65
FIGURA 6.2 - RELAÇÃO ENTRE AS INFORMAÇÕES QUE PERMITEM A MEDIÇÃO DOS VÍCIOS REFRATIVOS	66
FIGURA 6.3 - OPERAÇÕES EFETUADAS PELO MÓDULO DE PRÉ-PROCESSAMENTO DAS IMAGENS OFTALMOLÓGICAS.....	67
FIGURA 6.4 - IMAGEM DE HS APRESENTANDO OS ELEMENTOS DA IMAGEM REFERENCIADOS NA DESCRIÇÃO DOS ALGORITMOS DE PRÉ-PROCESSAMENTO	68
FIGURA 6.5 - EXEMPLOS DE IMAGENS DO TIPO “CLARA” E “ESCURA”	69
FIGURA 6.6 - ILUSTRAÇÃO DO CÁLCULO DO LIMIAR.....	70
FIGURA 6.7 - SEQUÊNCIA DE PASSOS DO PROCESSO DE CENTRALIZAÇÃO E PADRONIZAÇÃO DA IMAGEM	70
FIGURA 6.8 - IMAGEM COM RUÍDO ORIUNDO DO PROCESSO DE AQUISIÇÃO.	71
FIGURA 6.9 - SEQUÊNCIA DE PASSOS DO PROCESSO DE ELIMINAÇÃO DE RUÍDO.	72
FIGURA 6.10 - APLICAÇÃO DA PCA NAS COLUNAS DA IMAGEM PRÉ-PROCESSADA	74
FIGURA 6.11 - APLICAÇÃO DA PCA ÀS IMAGENS EM DUAS ETAPAS.....	75
FIGURA 6.12 - DIFERENÇA ENTRE OS VETORES DE CARACTERÍSTICAS OBTIDOS COM PCA – IMAGENS 1 E 2.....	77
FIGURA 6.13 - DIFERENÇA ENTRE OS VETORES DE CARACTERÍSTICAS OBTIDOS POR GABOR – IMAGENS 1 E 2.....	77
FIGURA 7.1 - MÓDULO DE ANÁLISE PROPOSTO INICIALMENTE COM MEDIDAS REGRESSIVAS. ...	85
FIGURA 7.2 - MÓDULO DE ANÁLISE PARA CLASSIFICAÇÃO.	88
FIGURA 8.1 - ARQUITETURA DO MÓDULO DE ANÁLISE UTILIZANDO COMBINAÇÃO DE CLASSIFICADORES.	102
FIGURA 8.2 - SUGESTÃO DE COMBINAÇÃO DE CLASSIFICADORES UTILIZANDO LÓGICA FUZZY.	102
FIGURA AN1.1 - DIFERENÇA ENTRE OS VETORES DE CARACTERÍSTICAS OBTIDOS COM PCA – IMAGENS 1 E 2.....	119
FIGURA AN1.2 - DIFERENÇA ENTRE OS VETORES DE CARACTERÍSTICAS OBTIDOS COM PCA – IMAGENS 2 E 3.....	120
FIGURA AN1.3 - DIFERENÇA ENTRE OS VETORES DE CARACTERÍSTICAS OBTIDOS COM PCA – IMAGENS 3 E 1.....	120
FIGURA AN1.4 - DIFERENÇA ENTRE OS VETORES DE CARACTERÍSTICAS OBTIDOS POR GABOR – IMAGENS 1 E 2.....	121

FIGURA AN1.5 - DIFERENÇA ENTRE OS VETORES DE CARACTERÍSTICAS OBTIDOS POR GABOR – IMAGENS 2 E 3.....	121
FIGURA AN1.6 - DIFERENÇA ENTRE OS VETORES DE CARACTERÍSTICAS OBTIDOS POR GABOR – IMAGENS 3 E 1.....	122
FIGURA AN2.1 - PORCENTAGENS DO NÚMERO DE PADRÕES E DO ERRO MÉDIO TOTAL PARA A MEDIDA S	124
FIGURA AN2.2 - PORCENTAGENS DO NÚMERO DE PADRÕES E DO ERRO MÉDIO TOTAL PARA A MEDIDA S COM PROPAGAÇÃO DO ERRO.	125
FIGURA AN2.3 - PORCENTAGEM DO NÚMERO DE PADRÕES E DO ERRO MÉDIO TOTAL PARA A MEDIDA C	126
FIGURA AN2.4 - PORCENTAGENS DO NÚMERO DE PADRÕES E DO ERRO MÉDIO TOTAL PARA A MEDIDA C COM PROPAGAÇÃO DO ERRO.....	126
FIGURA AN2.5 - PORCENTAGEM DO NÚMERO DE PADRÕES E DO ERRO MÉDIO TOTAL PARA A MEDIDA A	127
FIGURA AN2.6 - PORCENTAGENS DO NÚMERO DE PADRÕES E DO ERRO MÉDIO TOTAL PARA A MEDIDA A COM PROPAGAÇÃO DO ERRO.	128
FIGURA AN2.7 - PORCENTAGEM DO NÚMERO DE PADRÕES E DO ERRO MÉDIO TOTAL PARA A MEDIDA S COM BALANCEAMENTO ARTIFICIAL.....	129
FIGURA AN2.8 - PORCENTAGENS DO NÚMERO DE PADRÕES E DO ERRO MÉDIO TOTAL PARA A MEDIDA S COM BALANCEAMENTO ARTIFICIAL E PROPAGAÇÃO DO ERRO.....	129
FIGURA AN2.9 - PORCENTAGEM DO NÚMERO DE PADRÕES E DO ERRO MÉDIO TOTAL PARA A MEDIDA C COM BALANCEAMENTO ARTIFICIAL.....	130
FIGURA AN2.10 - PORCENTAGENS DO NÚMERO DE PADRÕES E DO ERRO MÉDIO TOTAL PARA A MEDIDA C COM BALANCEAMENTO ARTIFICIAL E PROPAGAÇÃO DO ERRO.....	130
FIGURA AN2.11 - PORCENTAGENS DO NÚMERO DE PADRÕES E DO ERRO MÉDIO TOTAL PARA A MEDIDA A COM BALANCEAMENTO ARTIFICIAL.....	131
FIGURA AN2.12 - PORCENTAGENS DO NÚMERO DE PADRÕES E DO ERRO MÉDIO TOTAL PARA A MEDIDA A COM BALANCEAMENTO ARTIFICIAL E PROPAGAÇÃO DO ERRO.	131

Lista de Tabelas

TABELA 3.1 - POLINÔMIO DE ZERNIKE ELEVADO AO QUARTO GRAU [LIA94].	34
TABELA 3.2 - MEDIÇÃO DO DEFOCUS E ATIGMATISMO DOS OLHOS DOS PACIENTES L E B [LIA94].	35
TABELA 3.3 - COEFICIENTES DE ZERNIKE PARA OS OLHOS DOS PACIENTES L E B [LIA94].	35
TABELA 5.1 – NÍVEIS DE ABSTRAÇÃO DOS DADOS DE ENTRADA PARA CADEIA DE PROCESSAMENTO DE IMAGENS [EGM02].	56
TABELA 5.2 - NÚMERO DE APLICAÇÕES EM QUE AS RNAs REALIZAM UMA TAREFA ESPECÍFICA DA CADEIA DO PROCESSAMENTO DE IMAGEM [EGM02].	57
TABELA 5.3 - DIFERENTES TIPOS DE REDES NEURAS APLICADAS ÀS VÁRIAS TAREFAS NA CADEIA DE PROCESSAMENTO DE IMAGEM [EGM02].	57
TABELA 6.1 –DISTÂNCIAS EUCLIDIANAS DOS VETORES DE CARACTERÍSTICAS OBTIDOS COM PCA..	74
TABELA 6.2 –DISTÂNCIAS EUCLIDIANAS DOS VETORES DE CARACTERÍSTICAS OBTIDOS POR GABOR ..	76
TABELA 7.1 – EXEMPLOS DA AMBIGÜIDADE DAS MEDIDAS DE S, C E A ASSOCIADAS À IMAGENS DISTINTAS DO MESMO OLHO..	80
TABELA 7.2 – QUANTIDADE DE EXEMPLOS DA MEDIDA S DO NOVO BANCO DE IMAGENS	81
TABELA 7.3 – QUANTIDADE DE EXEMPLOS DA MEDIDA C DO NOVO BANCO DE IMAGENS.....	81
TABELA 7.4 – QUANTIDADE DE EXEMPLOS DA MEDIDA A DO NOVO BANCO DE IMAGENS.....	82
TABELA 7.5 – QUANTIDADE DE EXEMPLOS DA MEDIDA A QUE SATISFAZEM O CRITÉRIO DE NÚMERO MÍNIMO.....	83
TABELA 7.6 – TIPOS DE KERNEL PARA MODELOS DE SVM [GUN98].....	84
TABELA 7.7 – DESEMPENHO DAS MELHORES CONFIGURAÇÕES DE SVM PARA A ARQUITETURA REGRESSIVA	86
TABELA 7.8 – DESEMPENHO DAS MELHORES CONFIGURAÇÕES DE SVM PARA A ARQUITETURA DE CLASSIFICAÇÃO.	87
TABELA 7.9 – RESULTADOS DA ARQUITETURA DE CLASSIFICAÇÃO COM 1400 ATRIBUTOS.	88
TABELA 7.10 – RESULTADOS DA ARQUITETURA DE CLASSIFICAÇÃO COM 14000 ATRIBUTOS. ...	89
TABELA 7.11 – RESULTADOS DA ARQUITETURA DE CLASSIFICAÇÃO COM 28000 ATRIBUTOS. ...	89
TABELA 7.12 – BALANCEAMENTO ARTIFICIAL DA MEDIDA S – 5 CLASSES (150 EXEMPLOS). ...	90
TABELA 7.13 – BALANCEAMENTO ARTIFICIAL DA MEDIDA C – 4 CLASSES (120 EXEMPLOS)....	90

TABELA 7.14 – BALANCEAMENTO ARTIFICIAL DA MEDIDA A – 4 CLASSES (100 EXEMPLOS)....	90
TABELA 7.15 – RESULTADOS DA ARQUITETURA DE CLASSIFICAÇÃO COM BALANCEAMENTO ARTIFICIAL (6400 ATRIBUTOS).	91
TABELA 7.16 – RESULTADOS DA ARQUITETURA DE CLASSIFICAÇÃO UTILIZANDO GABOR (200 ATRIBUTOS).	91
TABELA 7.17 – RESULTADOS DA ARQUITETURA DE CLASSIFICAÇÃO COM BALANCEAMENTO ARTIFICIAL UTILIZANDO GABOR.....	92
TABELA 7.18 – COMPARAÇÃO ESTATÍSTICA ENTRE OS ALGORITMOS PCA E GABOR.	93
TABELA 7.19 – COMPARAÇÃO ESTATÍSTICA ENTRE OS ALGORITMOS PCA E GABOR COM BALANCEAMENTO ARTIFICIAL.....	93
TABELA 7.20 – RESULTADOS DA ARQUITETURA DE CLASSIFICAÇÃO UTILIZANDO GABOR E REDES RBF.	95
TABELA 7.21 – RESULTADOS COM BALANCEAMENTO ARTIFICIAL UTILIZANDO GABOR E REDES RBF.....	95
TABELA 7.22 – RESULTADOS DA COMPARAÇÃO ESTATÍSTICA DOS ALGORITMOS SVM E RBF..	95
TABELA 7.23 – RESULTADOS DA COMPARAÇÃO ESTATÍSTICA DOS ALGORITMOS SVM E RBF COM BALANCEAMENTO ARTIFICIAL	96

Lista de termos e siglas

A	Eixo do astigmatismo
AG	Algoritmos Genéticos
AM	Aprendizado de Máquina
C	Cilíndrico (astigmatismo)
D	Dioptria
DE	Distância Euclidiana
ERM	<i>Empirical Risk Minimization</i>
HS	Hartmann-Shack
ICL	<i>Implant Contact Lens</i>
LASIK	<i>Laser-Assisted in Situ Keratomileusis</i>
MLP	<i>Multilayer Perceptron</i>
MNN	<i>Modular Neural Network</i>
MSE	<i>Mean Square Error</i>
PCA	<i>Principal Component Analysis</i>
PRK	<i>PhotoRefractive Keratoplasty</i>
RBF	<i>Radial Basis Function</i>
RNAs	Redes Neurais Artificiais
ROI	<i>Region Of Interest</i>
SH	Shack-Hartmann
SRM	<i>Structural Risk Minimization</i>
S	Esférico (miopia e hipermetropia)
SVMs	<i>Support Vector Machines</i>
Lenslet	Conjunto de lentes
Wavefront	Frente de onda

Resumo

VALERIO NETTO, A (2003). *Processamento e análise de imagens para medição de vícios de refração ocular*, Tese (Doutorado), Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos – SP, 142 Pp.

Este trabalho apresenta um sistema computacional que utiliza técnicas de Aprendizado de Máquina (AM) para auxiliar o diagnóstico oftalmológico. Trata-se de um sistema de medidas objetivas e automáticas dos principais vícios de refração ocular, astigmatismo, hipermetropia e miopia. O sistema funcional desenvolvido aplica técnicas convencionais de processamento a imagens do olho humano fornecidas por uma técnica de aquisição chamada Hartmann-Shack (HS), ou Shack-Hartmann (SH), com o objetivo de extrair e enquadrar a região de interesse e remover ruídos. Em seguida, vetores de características são extraídos dessas imagens pela técnica de transformada *wavelet* de Gabor e, posteriormente, analisados por técnicas de AM para diagnosticar os possíveis vícios refrativos presentes no globo ocular representado. Os resultados obtidos indicam a potencialidade dessa abordagem para a interpretação de imagens de HS de forma que, futuramente, outros problemas oculares possam ser detectados e medidos a partir dessas imagens. Além da implementação de uma nova abordagem para a medição dos vícios refrativos e da introdução de técnicas de AM na análise de imagens oftalmológicas, o trabalho contribui para a investigação da utilização de Máquinas de Vetores Suporte e Redes Neurais Artificiais em sistemas de Entendimento/Interpretação de Imagens (*Image Understanding*). O desenvolvimento deste sistema permite verificar criticamente a adequação e limitações dessas técnicas para a execução de tarefas no campo do Entendimento/Interpretação de Imagens em problemas reais.

Palavras-chave: Entendimento/Interpretação de Imagens, Sistemas Inteligentes, Aprendizado de Máquina, Máquinas de Vetores Suporte (SVM), Redes Neurais Artificiais, Optometria, Hartmann-Shack, Vícios Refrativos, Auto-refratores, Engenharia Biomédica.

Abstract

VALERIO NETTO, A (2003). *Image Processing and Analysis for Measuring Ocular Refraction Errors, Thesis (Doctorate), Institute of Mathematics and Computer Science, University of São Paulo, São Carlos – SP, 142 Pp.*

This work presents a computational system that uses Machine Learning (ML) techniques to assist in ophthalmological diagnosis. The system developed produces objective and automatic measures of ocular refraction errors, namely astigmatism, hypermetropia and myopia from functional images of the human eye acquired with a technique known as Hartmann-Shack (HS), or Shack-Hartmann (SH). Image processing techniques are applied to these images in order to remove noise and extract the regions of interest. The Gabor wavelet transform technique is applied to extract feature vectors from the images, which are then input to ML techniques that output a diagnosis of the refractive errors in the imaged eye globe. Results indicate that the proposed approach creates interesting possibilities for the interpretation of HS images, so that in the future other types of ocular diseases may be detected and measured from the same images. In addition to implementing a novel approach for measuring ocular refraction errors and introducing ML techniques for analyzing ophthalmological images, this work investigates the use of Artificial Neural Networks and Support Vector Machines (SVMs) for tasks in Image Understanding. The description of the process adopted for developing this system can help in critically verifying the suitability and limitations of such techniques for solving Image Understanding tasks in "real world" problems.

Keywords: *Image Understanding, Intelligent Systems, Machine Learning, Support Vector Machines (SVM), Artificial Neural Networks, Optometry, Hartmann-Shack, Refractive Errors, Auto-refractor, Biomedical Engineering.*

Capítulo 1

---

1. INTRODUÇÃO

Uma forma eficiente e relativamente barata de medir os erros de refração ocular (miopia, hipermetropia e astigmatismo), também conhecidos como vícios de refração, ou vícios refrativos, faz uso de um equipamento constituído de lentes-teste e figuras projetadas em um anteparo colocado a uma certa distância do paciente, que deve reconhecer perfeitamente as imagens projetadas. Em síntese, a rotina clínica de um exame de refração consiste de três passos fundamentais [Ven95]:

- Cicloplegiar o paciente, i.e., paralisar o músculo ciliar do olho, processo popularmente conhecido como “dilatar” a pupila;
- Por meio da retinoscopia, avaliar qualitativamente e quantitativamente o erro de refração ocular;
- Em seguida, submeter o paciente a um teste, em um equipamento denominado “Green’s”, cujo procedimento de medida utiliza optotipos constituídos por letras ou figuras projetadas em um anteparo distante 5m do paciente. Este deve reconhecer as letras ou figuras com o auxílio de várias lentes posicionadas em frente aos seus olhos.

Tal avaliação permite determinar as lentes corretivas mais adequadas ao paciente. Entretanto, a avaliação da lente corretiva é subjetiva, uma vez que é o paciente quem determina, em parte, a lente à qual melhor se adapta. No intuito de automatizar essas medidas de refração e diminuir o tempo de medição, que geralmente é em torno de 20 a 30 min, é muito útil contar com um equipamento automático e objetivo capaz de proporcionar a mesma, ou melhor, precisão do que as medidas tradicionais, em menor tempo. Isso favorece um aumento do número de exames feitos em pacientes no mesmo período de tempo de atendimento [Sno99].

Esse tipo de sistema, denominado refrator automático, ou auto-refrator, existe comercialmente, e é bem aceito por clínicas, consultórios e hospitais em que há uma demanda muito grande de pacientes [Can97] [Hum98] [Nik98] [Nid98] [Top98] [Ste99]. O auto-refrator é um equipamento que, além de não requerer muito espaço físico para a realização dos exames, é muito rápido em suas medidas, permitindo um aumento significativo no número de pacientes atendidos em um determinado período. Também é particularmente interessante para pessoas com muita sensibilidade à luz. Entretanto, a tecnologia para fabricação desses equipamentos está concentrada em poucas empresas estrangeiras e os projetos são todos patenteados, impossibilitando qualquer alteração ou adição de recursos por terceiros. Dessa forma, o desenvolvimento de equipamentos capazes de medir precisamente os erros refrativos oculares, de um modo objetivo e automático, em condições normais de visão dos pacientes, vem sendo encorajado por profissionais da área de oftalmologia [Ven96].

1.1 Objetivo e motivações

Buscou-se, neste projeto, investigar uma abordagem inédita para a medição de vícios refrativos (astigmatismo, hipermetropia e miopia) baseada em análise de imagens utilizando técnicas de Aprendizado de Máquina. O Aprendizado de Máquina (AM) é um campo de pesquisa da Inteligência Computacional cujo objetivo é o desenvolvimento de técnicas computacionais sobre o aprendizado, bem como a construção de sistemas capazes de adquirir conhecimento de forma automática. Um sistema de aprendizado é um programa de computador que toma decisões baseadas em experiências acumuladas por meio da solução bem sucedida de problemas anteriores [Mon03]. A motivação para a utilização das técnicas de AM neste projeto vem do grande progresso em pesquisa e aplicações práticas nos últimos anos que certificam a sua eficiência em extrair conceitos (conhecimento) a partir de amostras de dados [Mit97], particularmente no caso de Redes Neurais Artificiais (RNAs) [Rum86] [Mor91] [Mar92] [Vas95] [Cos96] [Car99] e Máquinas de Vetores Suporte (*Support Vector Machines* – SVMs) [Bal98] [Joa99].

Um dos maiores desafios para o projeto foi identificar uma técnica para aquisição das informações funcionais sobre o olho humano que permitisse o diagnóstico dos vícios refrativos. Imagens digitalizadas da íris ou da retina não contêm informações

funcionais para o diagnóstico. Aprofundando o estudo na área de Optometria, em busca de soluções, deparou-se com algumas técnicas de aquisição, até a descoberta da técnica Hartmann-Shack (HS) ou Shack-Hartmann (SH). Esta vem inovando as pesquisas na área de Optometria e, na opinião de especialistas, o medidor de anomalias (*aberrometer*) baseado no princípio de HS é o mais promissor dos vários métodos existentes [Thi99b]. A primeira aplicação desse princípio na medição de anomalias do olho humano ocorreu em 1994 [Lia94]. Desde então, vários pesquisadores têm se dedicado à melhoria da técnica, propondo modificações, adaptações e sugerindo novas formas de aplicação para outras condições clínicas [Hon01a]. Essa tecnologia de aquisição de dados do olho já é utilizada em equipamentos que representam a nova geração de instrumentos para auxiliar nas cirurgias refrativas a laser (*Photorefractive Keratoplasty* - PRK) [Pet99] [Gob99] [Sam01].

Este contexto motivou a decisão de usar imagens do olho adquiridas pela técnica de HS. Em vista da inviabilidade técnica e financeira de montar um protótipo do *hardware* de aquisição, buscou-se uma parceria com uma das instituições de pesquisa no exterior que, à época, possuía o *hardware* de aquisição. Foram realizados contatos com o grupo de Optometria da Universidade de Indiana (EUA), coordenado pelo Prof. Dr. Larry Thibos, que disponibilizou seu banco de imagens. O Professor Thibos demonstrou grande interesse no projeto, por vislumbrar a possibilidade da aplicação de técnicas de análise de imagens baseadas em algoritmos de AM também em imagens geradas por outras técnicas, como *Fluorescein*, *Retro-illumination* e *Optical scatter map* [Thi02a].

O reconhecimento e análise das características pertinentes ao globo ocular utilizando algoritmos de AM pode viabilizar não apenas o diagnóstico dos vícios de refração, como também a obtenção de outras informações sobre o olho a partir da mesma imagem de HS [Val01] [Val02]. Os sistemas de análise atuais baseados na técnica de HS utilizam um globo artificial emétrepe para calibrar o equipamento, isto é, dados são extraídos da imagem do globo artificial e comparados com os dados correspondentes extraídos das imagens obtidas do olho sob exame para verificar a presença de anomalias. Esse é o caso da técnica utilizada por Liang [Lia94]. Com isso, muitas informações contidas nas imagens são ignoradas no processo de análise [Him01] [Val01].

A proposta de utilizar técnicas de AM para a análise e interpretação de imagens de HS busca proporcionar um sistema de medição capaz de considerar o conteúdo da

imagem como um todo, e não apenas comparar dados discretos extraídos da mesma com dados de referência, como é feito hoje. A interpretação da imagem como um todo torna possível, em princípio, medir outras enfermidades do olho, como é o caso do “*Tears Film Breakup*¹” [Him01] o que é inviável utilizando as técnicas convencionais de análise, como os polinômios de Zernike [Lia94]. A abordagem adotada difere da que vem sendo usada pelos especialistas em Optometria/Oftalmologia para a medição de anomalias a partir de imagens [Val02], e oferece um leque de possibilidades em termos de pesquisa acadêmica e tecnológica voltada à interpretação de informações de um globo ocular. Até onde se tem conhecimento, este trabalho é pioneiro na utilização de técnicas de AM nas áreas de Optometria/Oftalmologia, podendo dar início a uma aproximação entre as áreas de Computação e Optometria.

1.2 Visão geral do sistema

O sistema de medição de vícios refrativos desenvolvido engloba quatro etapas distintas, apresentadas na Figura 1.1.

Figura 1.1 – Visão geral do sistema de medição dos vícios refrativos.

¹ O *tear film* é a primeira e a mais importante superfície de refração do olho; irregularidades na sua camada podem causar aberrações ópticas. Além disso, se o *tear film* se romper (*tear film breakup*) e expor a superfície áspera da córnea, o resultado será a dispersão da luz. Esses fatores diminuem a qualidade da imagem na retina e compromete a função visual [Him01].

- **Etapa 0:** Módulo de aquisição de imagens. A aquisição das imagens oftalmológicas pela técnica de Hartmann-Shack (HS) foi feita utilizando um equipamento construído pelo Grupo de Optometria da Universidade de Indiana (EUA). O teste do sistema foi feito utilizando um banco de imagens já disponível, cedido pelo professor responsável pelo Grupo.
- **Etapa 1:** Módulo de pré-processamento de imagens. As imagens oftalmológicas não são geradas em um padrão que permita a sua utilização imediata. É necessário padronizá-las e, posteriormente, eliminar ruídos que podem distorcer os resultados do processo de extração de características. Para isso foram implementados algoritmos específicos de processamento de imagens inspirados no método de análise de histograma e que utilizam informações espaciais e geométricas do domínio da aplicação [Gon92] [Son99].
- **Etapa 2:** Módulo de extração de características. O objetivo da extração é diminuir a quantidade de dados de entrada para o módulo de análise. Duas técnicas de dois métodos diferentes foram implementadas e testadas. Do método de análise multivariada foi implementada a técnica de análise de componentes principais (*Principal Component Analysis* – PCA) [Jon82]; do método de transformada de *wavelet* foi implementada a técnica da transformada de Gabor [Gab46].
- **Etapa 3:** Módulo de análise. É responsável pela interpretação dos atributos gerados pelo módulo de extração de características. Foram realizados experimentos com duas técnicas de AM, a saber, Máquinas de Vetores Suporte (*Support Vector Machines* – SVMs) e Redes Neurais Artificiais (RNAs), especificamente redes RBF.

O sistema foi desenvolvido em LINUX, adotando a filosofia do “*software* livre”. Os sistemas computacionais “livres” têm um papel importante no desenvolvimento de pesquisas, pois reduzem os custos de compra de *software* proprietário e/ou pagamento de licenças. A utilização de *software* de domínio público permite também que os resultados das pesquisas sejam mais facilmente distribuídos, o que viabiliza a troca de informação e tecnologia entre pesquisadores e centros de pesquisa.

1.3 Organização da tese

O Capítulo 2 introduz vários termos e noções básicas sobre o funcionamento do olho humano. Na Seção 2.1 são brevemente apresentadas as principais partes constituintes do olho, e na Seção 2.2 os chamados vícios refrativos, miopia, hipermetropia, astigmatismo e presbiopia. Na Seção 2.3 são revisados, em ordem cronológica, alguns processos existentes para a medição dos vícios refrativos para que o leitor tenha uma visão geral da evolução desses processos e suas limitações.

No Capítulo 3 é descrita a técnica de Hartmann-Shack (HS). Na Seção 3.1 são apresentados um breve histórico da introdução da técnica na área oftalmológica e a motivação para a sua utilização neste projeto. O seu princípio de funcionamento e as limitações técnicas são introduzidos na Seção 3.2. Em seguida, nas Seções 3.3 e 3.4, é descrito o sistema pioneiro para medição de anomalias do olho humano utilizando a técnica de HS desenvolvido por Liang na década de 90.

No Capítulo 4 são apresentadas as técnicas utilizadas para a extração das características relevantes das imagens do olho humano. O objetivo da extração é reduzir a quantidade de dados de entrada do módulo de análise. São descritas as duas técnicas de extração de características implementadas, sendo que na Seção 4.1 é introduzida a técnica da PCA, e na Seção 4.2 é introduzida a transformada *wavelet* de Gabor.

O Capítulo 5 expõe as técnicas investigadas para a implementação do módulo de análise de imagens. O capítulo está organizado em 4 seções. Na Seção 5.1 são introduzidas noções teóricas sobre as RNAs. Nas Seções 5.2 e 5.3 são discutidos, respectivamente, a sua aplicação em reconhecimento e análise de imagens e aspectos relacionados ao uso de Algoritmos Genéticos (AGs) associados às redes. Por fim, na Seção 5.4 é apresentada a fundamentação teórica sobre SVMs.

No Capítulo 6 são apresentados o banco de imagens e a implementação dos módulos de pré-processamento e extração de características do sistema. Na Seção 6.1 é relatado o processo de criação do banco de imagens oftalmológicas de HS utilizado para o teste dos sub-sistemas de processamento e análise. Em seguida, na Seção 6.2, é explicado como foi concebido e implementado o módulo de pré-processamento de imagens. Por fim, na Seção 6.3 é descrita a implementação do módulo de extração de características utilizando o algoritmo da PCA e, posteriormente, a transformada *wavelet* de Gabor.

O Capítulo 7 relata como foi concebido e implementado o módulo de análise do sistema. São apresentadas as várias alternativas para a sua implementação e os resultados obtidos com cada uma delas, até encontrar a mais adequada para o sistema de medição dos vícios refrativos. Na Seção 7.1 são discutidos as limitações do banco de imagens utilizado e o seu impacto no módulo de análise. Na Seção 7.2 são descritos os experimentos e seus resultados com os atributos extraídos utilizando PCA e, posteriormente, analisados pelo algoritmo SVM. Foram testadas várias configurações variando a quantidade de atributos de entrada para o módulo de análise. Na Seção 7.3, são descritos os experimentos realizados com atributos extraídos pela transformada *wavelet* de Gabor. O intuito foi verificar qual das duas técnicas de extração resultaria em um melhor desempenho do módulo de análise. A Seção 7.4 apresenta os resultados de uma abordagem alternativa utilizando os mesmos atributos fornecidos pela transformada de Gabor, porém aplicando redes RBF para a análise. O objetivo foi avaliar o desempenho das SVMs em relação a uma outra técnica de AM.

O Capítulo 8 está dividido em duas seções. Na Seção 8.1 são discutidas as principais contribuições e apresentadas algumas considerações sobre o trabalho desenvolvido e suas limitações. Na Seção 8.2 são apresentadas sugestões para trabalhos futuros.

Capítulo 2

---

2. O OLHO HUMANO E OS VÍCIOS REFRACTIVOS

O sentido da visão é responsável pela maior parte das informações que uma pessoa adquire, sendo que o processo de receber estímulos visuais (luz) e convertê-los em imagens coerentes é uma das tarefas mais custosas para o cérebro humano. Este capítulo apresenta, na Seção 2.1, uma breve revisão sobre as principais partes constituintes do olho humano, e na Seção 2.2 os chamados erros de refração, também conhecidos como, vícios de refração ou vícios refrativos (miopia, hipermetropia, astigmatismo e presbiopia) decorrentes da dificuldade de focalização da imagem sobre a retina. Na Seção 2.3 é apresentado um histórico da evolução dos métodos utilizados para a medição dos vícios refrativos. Inicia-se com um breve relato sobre a retinoscopia, passando pelos optômetros e, finalmente, chegando aos refratores automáticos, ou auto-refratores.

2.1 Estrutura e funcionamento do globo ocular

Quando uma pessoa olha para um objeto, raios de luz são refletidos do objeto na direção do olho. Os raios que penetram no sistema óptico do globo ocular sofrem um desvio, e produzem uma imagem invertida do objeto na retina. Dessa região, a imagem é transportada na forma de impulsos elétricos ao cérebro, onde é interpretada na posição correta.

A Figura 2.1 mostra as principais componentes de um globo ocular humano.

Figura 2.1 - Olho humano e seus principais componentes [Ana00].

O processo de convergência dos raios de luz que incidem no olho é iniciado na córnea. A córnea é um tecido transparente e avascular (que não possui vasos). Ela age como uma janela refringente e protetora, através da qual passam os raios de luz em direção à retina. Tem um poder refrativo equivalente a uma lente de +43 dioptrias² e possui uma espessura de 0.4 mm a 0.9 mm no pólo anterior, a qual aumenta em direção à periferia. Ela é composta por várias camadas: Epitélio, Lâmina Anterior, Estroma, Lâmina Posterior e Endotélio [Olh00].

A seguir, mais internamente no globo ocular, é encontrado a íris, que é responsável pelo controle da quantidade de luz que penetra no olho mediante o aumento ou diminuição de sua abertura central, a pupila (região central e escura da íris). O diâmetro da pupila determina a quantidade de luz que entra no olho, sendo que o seu tamanho varia de acordo com a quantidade de luz no ambiente.

Imediatamente atrás da íris está o cristalino (ou lente natural do olho), a segunda estrutura que participa do sistema de foco ótico do olho. Ela é incolor e quase completamente transparente, medindo cerca de 4 mm de espessura e 9 mm de diâmetro. Quando o cristalino está acomodado, seu poder de refração é máximo. Essa lente se desenvolve durante toda a vida do indivíduo, e suas dimensões crescem pela superposição de novas camadas sobre as mais antigas, formando uma estrutura estratificada como uma cebola.

² Unidade de medida do poder refrativo de uma lente; equivale ao poder refrativo de uma lente com distância focal de um metro. O número de dioptrias de uma lente corresponde ao inverso da distância focal em metros. Assim, uma lente de três dioptrias tem uma distância focal de 1/3 de metro [Mic00].

O grande espaço do globo ocular atrás do cristalino é ocupado por um material gelatinoso transparente, denominado humor vítreo ou corpo vítreo que tem a função de manter a coesão e a pressão do globo ocular. Uma característica importante é que seu índice de refração é praticamente igual ao índice de refração da córnea, não produzindo, portanto, desvio significativo do caminho da luz no olho. Devido a isso, o raio de luz focado pela córnea e pelo cristalino, o atravessa completamente em direção à porção mais posterior do olho, a retina.

A retina é um tecido do sistema nervoso. Ela percebe a luz e produz estímulos que são transmitidos para o cérebro pelo nervo óptico. Esses impulsos elétricos são integrados no córtex cerebral, produzindo a sensação visual [Olh99]. Macula é uma região pequena da retina que contém células especiais sensíveis à luz. Ela é responsável pela percepção de detalhes finos da visão.

2.2 Erros de refração

A refração é a forma como o olho direciona a luz para focá-la na retina. A córnea e o cristalino são as lentes que refratam a luz, ou seja, direcionam os raios de luz em direção à retina. Essa capta a imagem formada pelos raios e transmite-a até o cérebro, por meio do nervo óptico. Quando o globo ocular apresenta alguma dificuldade de focalização da imagem sobre a retina que impede que a luz seja refletida adequadamente formando imagens alteradas, essa condição é genericamente definida como ametropia, ou erro de refração (vício refrativo), os quais podem ser de quatro tipos: miopia, hipermetropia, astigmatismo e presbiopia [Doe99] [Err99] [Glo99]. Os principais sintomas dos erros refrativos são a diminuição da visão, o desconforto nos olhos e, ocasionalmente, dores de cabeça. O uso de lentes de contato ou de óculos para compensar as dificuldades apresentadas normalmente é suficiente para eliminar os sintomas.

A miopia é um distúrbio de focalização da imagem pelo qual esta se forma anteriormente à retina. Isso se deve a um maior comprimento do globo ocular, ou aumento na curvatura da córnea ou cristalino, resultando em dificuldades para enxergar à distância. A miopia pode ser corrigida com lentes esféricas divergentes (Figura 2.2). O olho míope é mais comprido do que o normal e, conseqüentemente a luz tem uma distância maior a percorrer. A focalização não ocorre na retina como seria correto, mas

antes dela, provocando um embaçamento da imagem dependendo da distância do objeto ao observador (objetos próximos são vistos claramente).

Exames preventivos são recomendados para identificar a presença de miopia em crianças e antecipar a correção visual. Independentemente da idade, as pessoas míopes são mais propensas a apresentar um descolamento de retina, cuja única forma de tratamento é a cirurgia, o que indica a necessidade de exames preventivos regulares também nos adultos. Além das lentes de contato e dos óculos, em alguns casos pode ser indicado um procedimento cirúrgico para compensar esse erro. Entretanto, em determinados casos lentes corretivas podem ser necessárias mesmo após a cirurgia. Existem atualmente várias técnicas cirúrgicas para correção de miopia, sendo que duas técnicas bastante empregadas são [Tec99]: PRK (*PhotoRefractive Keratoplasty*) ou fotoablação superficial, no qual um *laser* é aplicado superficialmente sobre a córnea; e o LASIK (*Laser-Assisted in Situ Keratomileusis*) ou fotoablação conjugada, uma técnica mais avançada e com melhores características de cicatrização.

Figura 2.2 – Como uma lente pode ser utilizada para corrigir o erro refrativo causado pela miopia [Doe99].

Na hipermetropia também ocorre uma dificuldade de focalizar a imagem, mas contrariamente à miopia a focalização se dá posteriormente à retina. Ela se deve, portanto, a um globo ocular com menor comprimento, ou devido à córnea ou cristalino possuírem uma menor curvatura, e pode ser corrigida com lentes esféricas convergentes (Figura 2.3). A hipermetropia é bastante comum em crianças e pode vir a desaparecer com o crescimento dos olhos. Às vezes ela ocorre associada ao estrabismo, uma vez que os músculos têm que se contrair para ver de perto [Err99]. Existem várias técnicas

cirúrgicas para a correção da hipermetropia, como Rádio Frequência, PRK, LASIK ou ICL (*Implant Contact Lens*) [Tec99].

Figura 2.3 - Como uma lente pode ser utilizada para corrigir o erro refrativo causado pela hipermetropia [Doe99].

O astigmatismo é uma condição que decorre da diferença de curvatura da córnea ou cristalino nas direções horizontal e vertical (comparável às curvaturas de um ovo ou de uma bola de futebol americano). Disso resultam diferentes profundidades de foco que distorcem a visão, tanto de longe quanto perto. Ele pode ocorrer associado à miopia ou à hipermetropia e pode ser corrigido com lentes cilíndricas. Grande parte dos casos de astigmatismo é congênito e não se modifica ao longo da vida do indivíduo. Além dos óculos e das lentes de contato, a cirurgia também pode ser uma alternativa.

A presbiopia, ou vista cansada, é uma condição inevitável que surge, normalmente, em indivíduos após os 40 anos de idade, em média, e decorre de uma perda progressiva da flexibilidade do cristalino. Ela provoca dificuldade para a visão de perto, conhecida popularmente como vista cansada, e pode ser corrigida com lentes esféricas convergentes.

2.3 Processos de medição dos vícios refrativos

Os processos de medição para detectar erros de refração foram evoluindo a medida em que surgiam as necessidades de diminuir o tempo de atendimento aos pacientes, eliminar a subjetividade na avaliação da lente corretiva e melhorar a precisão das

medidas. Um dos primeiros equipamentos desenvolvidos para essa finalidade foi um retinoscópio, criado por Jack Copeland em 1927. Copeland desenvolveu uma técnica clínica importante e os instrumentos para aplicá-la rápida e facilmente [Slo70].

2.3.1 Retinoscopia

A retinoscopia pode ser didaticamente dividida em dois sistemas distintos: iluminação e observação. O primeiro refere-se apenas à incidência da luz no olho do paciente, e o segundo à luz retro-espalhada pelo fundo do olho desse paciente até atingir o olho do examinador, através do orifício central (Figura 2.4).

Figura 2.4 - Luz proveniente da retina do paciente passando pelo orifício central do espelho e atingindo a retina do examinador [Ven95].

Na prática, o examinador projeta uma imagem do fecho de luz no plano da pupila do paciente. Ele não vê diretamente a retina iluminada do paciente, mas a imagem ampliada pelo sistema ótico do olho observado. A luz projetada atravessa o olho observado, e é influenciada pelos seus componentes óticos. A maneira pela qual a luz é influenciada proporciona uma medida qualitativa do sistema ótico ocular do paciente. Numa situação hipotética, em que o examinador encontra-se no infinito, com um pequeno movimento lateral seqüencial da luz incidente (perpendicular) é possível avaliar qualitativamente a ametropia apresentada pelo paciente do seguinte modo (Figura 2.5):

Figura 2.5 - Esquematisação do modo como o retinoscopista verá o “reflexo” da luz na retina se sobrepondo à faixa de iluminação, em termos da dinâmica do movimento [Ven95].

- a) **Olho emétrepe**: (Figura 2.5a) embora haja um movimento da luz incidente, a luz refletida permanece estática, porque a imagem da faixa de luz está exatamente no foco do sistema ótico (retina). Diz-se, assim, que foi atingido o ponto de neutralidade;
- b) **Olho hipermétrepe**: (Figura 2.5b) a luz refletida na forma de uma faixa divergente movimenta-se no mesmo sentido da luz incidente (“movimento a favor”), porque o foco do sistema ótico ocular do paciente está posicionado atrás da retina;
- c) **Olho míope**: (Figura 2.5c) a luz refletida na forma de uma faixa convergente movimenta-se no sentido oposto ao da luz incidente (“movimento contra”), porque o foco do sistema ótico do paciente está posicionado anteriormente à retina [Ven95].

2.3.2 Optômetros

Os optômetros são um meio termo entre os retinoscópios e os refratores automáticos. Nas décadas de 20 e 30 vários optômetros foram desenvolvidos e comercializados. Esses equipamentos, apesar de classificados como instrumentos objetivos de medidas de refração, são na verdade subjetivos no que se refere ao examinador, que deve focalizar ou alinhar a imagem de um alvo na retina do paciente. Tais instrumentos são baseados em dois princípios de medição: o princípio de Scheiner e o do Optômetro, os quais têm sido amplamente utilizados em tentativas de se automatizar a refração clínica [Ven95].

O Princípio de Scheiner foi o primeiro de uma classe de métodos de avaliação de zona focal para sistemas de refração e de testes de componentes ópticos. Esses métodos têm sido amplamente empregados na tentativa de automatizar as medidas de refração ocular e a maioria dos refratores automáticos é baseada nesse princípio. Em 1619, Scheiner [Guy81] descobriu que o ponto em que o olho é focalizado pode ser precisamente determinado apenas colocando uma fenda dupla em frente à pupila. A luz que entra no olho proveniente de um objeto pontual posicionado no infinito é limitada pela fenda. Os raios se cruzam antes de alcançarem a retina se o olho for míope, atrás da retina se for hipermetrópe, e exatamente na retina se for normal. Assim, quando o olho é míope ou hipermetrópe, tem-se a presença de dois pontos na retina, enquanto que no olho normal tem-se apenas um (Figura 2.6). Movendo o objeto até a posição em que aparece um único ponto na retina pode-se determinar o ponto remoto do olho e, portanto, a correção refrativa. O erro refrativo do olho é igual ao inverso da distância do ponto remoto em metros.

Figura 2.6 - Esquematização do Princípio de Scheiner [Ven95].

O princípio do optômetro, descrito em 1759, permite uma variação contínua de potência em instrumentos de refração. Uma única lente convergente é utilizada e posicionada a uma certa distância do olho, ou do plano das lentes corretoras, igual à sua distância focal. Um feixe de luz a partir de um alvo do lado oposto da lente entra no olho com vergências diferentes (zero, positivas ou negativas), dependendo da posição do alvo. A vergência da luz no plano focal do optômetro é medida, observa-se que está linearmente relacionada com o deslocamento do alvo. Portanto, pode-se construir uma

escala com espaçamentos iguais que indicará o número de dioptrias que este arranjo ótico simula no plano das lentes corretoras.

Três fatores limitam a aceitação dos optômetros na prática clínica: problemas de alinhamento, astigmatismo irregular e acomodação. Alguns destes fatores, discutidos a seguir, ainda hoje são limitantes para o desempenho dos auto-refratores [Ven95].

- **Alinhamento:** Num sistema baseado no princípio de Scheiner, ambas as fendas devem encaixar exatamente na pupila de entrada do olho do paciente. Caso o paciente se mova ligeiramente, a medida é invalidada. Adicionalmente, alinhar o instrumento e manter o alinhamento requer uma grande habilidade prática e paciência por parte do examinador, bem como cooperação do paciente.
- **Astigmatismo Irregular:** Um sistema de Scheiner usa apenas duas pequenas porções do sistema ótico ocular. Na presença de astigmatismo irregular, a melhor refração média sobre toda a pupila pode ser completamente diferente da apresentada por apenas duas pequenas áreas da mesma. O melhor meio de evitar este problema é utilizar a maior área pupilar possível.
- **Acomodação:** Quando os pacientes olham dentro dos instrumentos, seus olhos tendem a acomodar. Essa situação, conhecida como miopia de instrumento ou acomodação de instrumento, é o maior problema para os optômetros, bem como para a maioria dos refratores automáticos. A quantidade de acomodação introduzida pelos equipamentos geralmente flutua durante o processo de medida. Se os meridianos principais são medidos sucessivamente, ao invés de simultaneamente, as flutuações da acomodação dos pacientes podem induzir um erro nas correções cilíndricas, bem como nas esféricas. Muitos fatores influenciam a acomodação, como atenção, fadiga, direção de fixação do olhar, iluminação, detalhes de imagem, desfocalização da imagem na retina e fatores psicológicos.

Os optômetros automáticos no infravermelho, que fazem as medidas de refração automaticamente utilizando luz infravermelha, que é invisível ao paciente, foram os que despertaram maior interesse. Um alvo visível é incorporado em cada instrumento para ajudar no controle da fixação do paciente e da acomodação. O primeiro optômetro dessa categoria foi descrito por Collins em 1937 [Cha76]; porém, a eletrônica da época não era suficientemente sofisticada para tornar o instrumento prático. Com o aprimoramento

dos fotodetectores de infravermelho, dos circuitos de controle e dos microprocessadores e, posteriormente, com o advento do laser, vários desses sistemas surgiram e originaram os sistemas de refração automáticos encontrados atualmente no mercado.

2.3.3 Refratores automáticos

A auto-refração é um teste automático que faz uma medição computadorizada do olho humano para determinar aproximadamente a prescrição das lentes para os óculos de um determinado paciente [Ste99]. O equipamento que faz esse teste é chamado de auto-refrator (Figura 2.7), e é frequentemente combinado com um ceratômetro (*Keratometer*), equipamento utilizado para medir a curvatura da córnea [Can99] [Ste99], fornecendo ao médico rapidez e precisão em ambas as medidas.

O médico utiliza os resultados do equipamento como ponto de partida para prescrever, por exemplo, um primeiro par de óculos para um paciente recente ou para determinar uma possível troca de lentes para um paciente que já usa óculos. Os auto-refratores também possibilitam ao médico fazer uma medição com razoável precisão em pacientes que não falam ou não são cooperativos, como indivíduos que sofrem de problemas mentais ou pacientes muito jovens.

Um estudo [Ura95] foi realizado por profissionais de oftalmologia para comparar a refração automática objetiva, que utiliza os auto-refratores, com a refração clínica ou subjetiva, sem a utilização do equipamento. A conclusão foi que a média das diferenças entre o aparelho e a refração subjetiva fica abaixo de 0,5 D (Dioptrias) para as medidas esféricas e cilíndricas; já com relação ao eixo do astigmatismo a diferença foi significativa, acima de 10° . Concluiu-se que os auto-refratores foram mais precisos nas medidas esféricas e cilíndricas, deixando a desejar na obtenção do eixo do astigmatismo. No entanto, pelo fato das medidas esféricas e cilíndricas serem confiáveis, os auto-refratores podem reduzir a duração do exame refratométrico. Isso é de grande valia especialmente no caso de crianças, pois diminui o cansaço, além de orientar o oftalmologista nos casos de má informação e não cooperação [San98].

Figura 2.7 – Vistas de um auto-refrator comercial [Can97] [Can99].

Na área acadêmica, Ventura [Ven93] [Ven94a] [Ven94b] [Ven96] pesquisou e analisou alguns sistemas de medição dos vícios refrativos e, fruto desse trabalho, propôs um sistema chamado DVL (Detecção de Vergência da Luz) em duas versões, DVL-1 e DVL-2 [Ven95]. De um modo geral, a automação de sistemas oftalmológicos para diagnóstico segue uma rotina em que o examinador é substituído por um detector, uma interface detector-microcomputador, e um *software* dedicado; e as fontes de luz, que geralmente são lâmpadas ou *leds*, por lasers. A evolução dos sistemas de medidas de vícios de refração não seguiu essa regra. Os refratores oculares manuais, os optômetros manuais e os automáticos, não projetam uma fonte luminosa no fundo do olho, que é o princípio básico da retinoscopia. Os sistemas DVL se baseiam exatamente numa automação da retinoscopia, ou seja, uma fonte luminosa é projetada na retina e a vergência da luz retro-espalhada pela mesma, e mostraram-se mais eficientes comparados a outros sistemas baseados na automação dos optômetros e refratores manuais. Os resultados obtidos com o DVL, tanto para as réplicas oculares como para olhos *in vivo*, são extremamente satisfatórios. Porém, o sistema é eficaz apenas para pacientes cicloplegiados. A cicloplegia ocular consiste em “paralisar” os músculos ciliares, responsáveis pela ação da acomodação visual, por meio da aplicação de colírios, como a atropina. A cicloplegia leva em torno de 45 minutos para ser atingida e a sua ação dura em média de 2 a 18h e, em certos casos, até sete dias [Ven95].

3. AQUISIÇÃO DE DADOS DO OLHO HUMANO

Existem vários métodos para a obtenção de informações provenientes do olho humano, sendo o medidor de anomalia (*aberrometer*) baseado no princípio de Hartmann-Shack (HS) ou Shack-Hartmann (SH) um dos mais promissores [Thi99b]. Neste capítulo são apresentadas informações relevantes sobre essa técnica, a qual foi utilizada para gerar as imagens analisadas pelo sistema desenvolvido neste projeto. Na Seção 3.1 é apresentado um breve histórico da introdução da técnica de HS na área oftalmológica e discutidos os motivos para a utilização de imagens de HS neste projeto. Na Seção 3.2 é apresentada uma breve revisão sobre o princípio de funcionamento da técnica e suas limitações práticas. Em seguida, nas Seções 3.3 e 3.4, é descrito o sistema pioneiro, desenvolvido por Liang na década de 90, para a medição de anomalias do olho humano utilizando a técnica de HS.

3.1 A técnica de Hartmann-Shack na área oftalmológica

Em 1900, Hartmann inventou um método para medir as anomalias dos raios de luz provenientes dos espelhos e lentes utilizados em telescópios. Ele usava um disco de metal com orifícios que podiam ser atravessados por raios de luz isolados. Os raios que seguiam um caminho errado eram chamados de raios com deformidade. Esse sistema ficou conhecido como “projektor de Hartmann”. Setenta anos mais tarde, em 1970, Roland Shack e Ben Platt [Sha71] inventaram um novo tipo de “projektor de Hartmann” feito com um arranjo de minúsculas lentes, que foi denominado sensor de Shack-Hartmann (SH). Este sensor foi utilizado, inicialmente, em projetos de laser para a aeronáutica dos EUA. Posteriormente, passou a ser utilizado por organizações que

atuam na área de ajustes e medições ópticas (telescópios, *lasers* e sistemas ópticos) [Wav00].

As primeiras pesquisas acadêmicas utilizando sensores de SH visavam a melhoria de imagens astronômicas e observações espaciais (estrelas, observações solares, etc.) [Har78] [Löf96] [Bar96] [Day00]. Mais tarde, pesquisadores começaram a utilizar esses sensores em sistemas de análise do globo ocular. Pode-se citar trabalhos como os de Liang et al. [Lia91], que objetivava a melhoria da resolução de medições ópticas dos aparelhos de LST (*Laser Tomographic Scanner*); de Miller et al. [Mil96], que trabalharam com imagens de cones foto receptores; de Walsh et al. [Wal84]; William et al. [Wil96] e Woods et al. [Woo96], que estudaram a utilização do sistema para a medição do coma, de anomalias esféricas e dos eixos cilíndricos.

Em 1994, Liang e colegas começam a utilizar o sensor de SH em sistemas de medição de anomalias do globo ocular [Lia94] [Lia97a]. No artigo de 1994, Liang refere-se à técnica como método Hartmann-Shack, o que gerou inúmeras argumentações sobre se o método deveria ser chamado Shack-Hartmann ou Hartmann-Shack. Hoje, os pesquisadores da área aceitam as duas denominações. Depois de Liang, outros pesquisadores publicaram artigos explorando a aplicação do sensor de HS na medição de anomalias oculares, denominando o procedimento de técnica de Hartmann-Shack para medição de frente de onda ou, simplesmente, técnica de Hartmann-Shack, ou método Hartmann-Shack [Kin97] [Ham97]. Em 1999, Thibos e Applegate lideraram uma “força tarefa” para recomendar metodologias e padrões para a divulgação das características das anomalias do olho humano [App99]. Em 2000 ocorreu o primeiro *International Congress of Wavefront Sensing and Aberration-free Refractive Correction* em Santa Fé, Novo México (EUA), o segundo foi realizado em 2001 em Monterey, California (EUA) e o terceiro em 2002 em Interlaken, Suíça. Os eventos buscavam fomentar a discussão de temas relacionados à aquisição das anomalias do olho humano, processamento e posterior análise das imagens resultantes, inclusive com envolvendo desde a opção de utilização dos descritores de Zernike, Taylor ou Seidel para a análise, até definições sobre a calibração do equipamento de medição [Thi00b].

Cientistas têm se dedicado a examinar, por exemplo, a replicabilidade e a precisão das medidas de anomalias obtidas pela técnica, bem como comparar tecnicamente esse sistema de medida com outros que têm a mesma finalidade, com resultados satisfatórios [Lia97a] [Lia97b] [Hon01]. Ash [Ash98] investigou o desempenho do sensor de HS

operando sob níveis baixos de luminosidade. Hamam [Ham01] propôs um método rápido para analisar padrões de HS que permite estudar as limitações do sistema e melhorar a capacidade de medir as aberrações. Vários trabalhos [Law96] [Rio97] [Sal99] na literatura propõem modificações na componente de aquisição da frente de onda nos sistemas que utilizam o sensor de HS. Thibos & Hong [Thi99b] verificaram a eficiência da técnica de HS para a medição de anomalias do olho para quatro classes de condições clínicas: anomalias para *dry eye*, enfermidade na córnea (Keratocomus), cirurgia refrativa na córnea (*laser-assisted in situ keratomileusis* – LASIK), e catarata. Eles concluíram que, em cada um desses casos, é possível obter no mínimo um mapa parcial topográfico da refração das anomalias do olho do paciente, porém com possibilidade de ocorrerem graves perdas na integridade dos dados [Thi99b].

Um trabalho pioneiro envolvendo pesquisadores da área de Computação Gráfica do Departamento de Computação e pesquisadores da Escola de Optometria, ambos da Universidade da Califórnia, Berkeley, foi concluído em 2001 [Opt02]. Trata-se de um sistema para “*rendering* realístico da visão” (RAYS – *Render As You See*) que simula a visão real dos pacientes. Os dados dos pacientes são fornecidos por um sistema de HS. Entrando com a distância focal, o algoritmo do RAYS apresenta a imagem com as distorções correspondentes à visão do paciente [Bar01].

Apesar da sua concepção ser simples, a técnica de HS só foi validada recentemente para a medição dos olhos. Isso decorre das dificuldades de produzir um equipamento de aquisição de frente de ondas (*wavefront*) utilizando o sensor de HS [Oko00] adaptado para a medição de globos oculares, principalmente no meio de pesquisa acadêmico [Thi99a] [Thi99b]. Porém, diversas universidades possuem protótipos operacionais do equipamento, como é o caso da *School of Optometry* da *Indiana University* (EUA) [Vis01], da *School of Optometry* da *University of Montreal* (Canadá) [Ham01], do *Centre for Eye Reseach* da *University of Melbourne* (Austrália) [Cen01], e da *University of Heidelberg* (Alemanha) [Fas01].

No meio industrial também existem equipamentos baseados na técnica [Fun01]. É o caso do *Wavefront Sciences Optics and Instrumentation* com o “*Complete Ophthalmic Analysis System*™” (COAS™), por sinal a empresa pioneira na área [Wav00]; do “*WaveScan*™ *Wavefront System*” da VISX [Wav01]; do “*CustomCornea Wavefront device*” da *Alcon Summit Autonomous* [Fun01]; e do “*Surgical Technolas 217 Laser*” da empresa Bausch & Lomb [Fun01], um sistema de *wavefront* acoplado a um

equipamento para operações a *laser*. Além desses, quase todas as grandes multinacionais da área de equipamentos oftalmológicos (Canon, Humphrey, Nikon, Nidek, Topcon) estão viabilizando sua versão de aparelho utilizando o método HS. Essa forma de aquisição vem sendo utilizada em equipamentos que representam a nova geração de instrumentos para auxiliar nas cirurgias refrativas a *laser* (*Photorefractive Keratoplasty* - PRK) [Pet99] [Gob99] [Sam01]. Os dados sobre as anomalias oculares, juntamente com a topografia de córnea [Car98], podem servir como entrada para o computador que controla o *laser*. Isso é possível porque a topografia de córnea informa as elevações em vários pontos da córnea antes da cirurgia. Algoritmos que relacionam os dados da topografia com aqueles das anomalias permitem informar ao *laser* o perfil exato do olho para a cirurgia [Kle98].

3.2 Princípio de funcionamento

A seguir são apresentadas duas descrições da técnica de HS, primeiro usando a terminologia dos raios ópticos (*ray optics*) e, posteriormente, utilizando a terminologia de ondas ópticas (*wave optics*). Busca-se, assim, oferecer uma visão contextual do método segundo as duas abordagens mais utilizadas na Optometria [Thi00a].

Há 400 anos atrás, em 1619, um célebre jesuíta e filósofo chamado Scheiner, contemporâneo a Kepler e Galileu, criou um dispositivo simples para demonstrar o mecanismo de focalização dos olhos, conhecido como o disco de Scheiner. Como mostrado na Figura 3.1, se um olho defeituoso (amétrope) observa um ponto de luz a uma certa distância, através de um disco que possui dois orifícios (*pinholes*), serão formadas duas imagens na retina. Se a imperfeição do olho é um caso simples de *defocus* (miopia ou hipermetropia) ou astigmatismo, então as imagens duplas da retina podem ser corrigidas por uma lente de grau apropriado.

Figura 3.1 – Princípio do disco de Scheiner [Thi00a].

Figura 3.2 – Optômetro que utiliza o princípio de Scheiner [Thi00a].

Observando a Figura 3.2, se os raios de luz que atravessam os dois orifícios do disco de Scheiner vêm de fontes diferentes, uma fonte pode permanecer fixa enquanto a outra pode ser reposicionada. Ajustando o movimento da fonte na horizontal e vertical o raio de luz é redirecionado até que se cruze com o raio fixo, gerando um único ponto de luz na retina. Feito esse ajuste, os deslocamentos Δx e Δy correspondem à imperfeição do raio de luz proveniente da pupila. Uma versão computadorizada desse conceito básico aplicado em um dispositivo que recebeu o nome de “*spatially resolved refractometer*” foi descrito por Webb et al. em 1992 [Web92].

Para converter o método subjetivo de Scheiner em um optômetro objetivo a direção da propagação da luz é invertida: um ponto de luz é colocado na retina, o qual se torna um ponto origem que reflete a luz proveniente do olho. São feitos mais alguns orifícios no disco de Scheiner tornando-o, assim, um Hartmann *Screen* (Figura 3.3). Cada uma dessas aberturas no Hartmann *Screen* isola um feixe fino do raio de luz que emerge do olho vindo de uma região específica da pupila. Os raios então interceptam um sensor

CCD (*Charge-Coupled Device*) [Sal81], que mostra a posição horizontal e vertical dos mesmos. Esse é o princípio do equipamento denominado “*aberrometer*” (medidor de anomalias) de Hartmann. Preenchendo cada um dos orifícios individuais do Hartmann *Screen* com lentes minúsculas, tem-se o *aberrometer* de Hartmann-Shack, ou um *aberrometer* de Scheiner-Hartmann-Shack [Sha71].

Figura 3.3 – Visão geral do *aberrometer* de Hartmann [Thi00a].

A seguir, a técnica é descrita utilizando a terminologia de ondas ópticas (*wave optics*). O propósito do *aberrometer* de Scheiner-Hartmann-Shack é medir uma função de onda do sistema óptico de um olho, o que é equivalente a medir a forma da frente de onda de luz refletida para fora do olho proveniente de um ponto de origem no *fundus* (fundo do olho). Por exemplo, uma frente de onda de luz emitida para fora de um olho perfeito deveria ser uma parte circular da onda plana com o mesmo diâmetro da pupila. Como o *aberrometer* de HS mede formas de frente de onda, pode-se dizer que se trata de um tipo de *fundus camera* [Thi00a]. Ao contrário de um *fundus camera* convencional, que focaliza toda a luz refletida de um ponto da retina para dentro de uma imagem simples (Figura 3.4), o *aberrometer* tem um conjunto de lentes minúsculas, semelhante ao olho de um inseto que divide a luz em vários feixes individuais, produzindo múltiplas imagens do mesmo ponto de luz proveniente da retina (Figura 3.5).

Figura 3.4 – Descrição de *fundus camera* convencional [Thi00a].

Figura 3.5 – Descrição de um *fundus camera* modificado [Thi00a].

Na prática, para a medição das anomalias do olho, um ponto origem de luz é criado na retina, e a frente de onda emergente desse olho é analisada por um conjunto de microlente (*lenslet*) que divide essa frente de onda em vários feixes individuais, cada qual focado no sensor CCD. Para um olho perfeito, uma onda plana refletida será focada perfeitamente dentro de uma grade de imagens de pontos (Figura 3.6). Entretanto, se o olho possuir alguma anomalia, a frente de onda será distorcida. A Figura 3.7 mostra a reflexão de uma frente de onda distorcida proveniente de um olho amétrope. Os raios de luz refletidos em um olho imperfeito não são paralelos, conseqüentemente, ao atingir o conjunto de micro lentes as imagens dos pontos serão focadas desordenadamente. Uma análise da localização dos pontos registrados no sensor CCD permite calcular matematicamente a inclinação (*slope*) da frente de onda com anomalia. Essa inclinação é fundamental para a qualidade óptica dos olhos, sendo chamada de “função da imperfeição da frente de onda” (*wavefront aberration function*). Essa função é parte principal do “coração” de uma teoria óptica que permite calcular a imagem formada na retina de algum objeto, para avaliar a qualidade dessa imagem e diagnosticar o

desempenho visual em tarefas que necessitem do olho para serem realizadas. Entretanto, para aplicar essa teoria óptica é necessário analisar a frente de onda imediatamente após ela passar pela pupila. Para isso, é necessário usar um par de lentes *relay* que focam o conjunto de micro lentes dentro da pupila (Figura 3.8). Essa configuração final é a forma básica do *aberrometer* de Scheiner-Hartmann-Shack ou Hartmann-Shack, desenvolvido por Liang e seus colegas em 1994 [Lia94].

Figura 3.6 - Princípio da análise da frente de onda em um olho emétopo [Thi99a].

Figura 3.7 - Princípio da análise da frente de onda em um olho amétopo [Thi99a].

Figura 3.8 - Configuração final do “*aberrometer*” Hartmann-Shack [Thi00a].

Existem, basicamente, duas limitações da técnica HS que Thibos [Thi00a] denomina de escala pequena e escala grande das anomalias. A suposição básica da técnica é que as microlentes têm diâmetros pequenos comparados às distorções presentes em uma frente de onda óptica. Essa suposição começa a falhar quando a magnitude das anomalias é grande (anomalias de baixa ordem). Nesse caso, a frente de onda “curva-se” sobre as aberturas das microlentes e o resultado é um ponto borrado que é dificilmente localizado. Se as anomalias são suficientemente grandes, os pontos podem se sobrepor, o que dificulta consideravelmente a análise da imagem resultante. A segunda limitação está presente em anomalias que possuem uma escala muito pequena. Essas microanomalias, também chamadas de anomalias de alta ordem, dispersam a luz e borram os pontos formados pelo sistema. Apesar desses pontos borrados serem problemáticos, eles contêm informações importantes sobre o grau e a localização da origem da dispersão dentro do olho.

3.3 Sistema de aquisição

Na Figura 3.9 é apresentado o diagrama esquemático da configuração do sistema de aquisição que Liang e colegas [Lia91] [Lia94] utilizaram para medir as anomalias do olho. O sistema consistia de um *laser* He-Ne com a saída de 1.5 mW a 632.8 nm, um sensor de frente de onda de HS formado por um conjunto de 15 x 15 micro lentes, um módulo de imagem CCD (*Charge-Coupled Device*) [Sal81], um computador (PC) com uma placa de aquisição de imagem para digitalizar os sinais de vídeo (512 *pixels* X 512 *pixels*, 8 *bits*), um *shutter* eletrônico, e componentes ópticos.

Primeiramente o *laser* é atenuado por um filtro de densidade neutra (*Neutral Density Filter – NDF*), que diminui a intensidade de luz que penetra no olho para um nível de $5\mu\text{ W}$. Posteriormente, o feixe de laser passa por um *shutter* controlado por um computador (PC), e é filtrado e ampliado para uma onda plana pelo colimador (3 na Figura). Uma íris (4) é usada para controlar o tamanho iluminado da pupila. O tamanho da pupila do olho iluminado para o experimento realizado foi de, aproximadamente, 3 mm. Em seguida, o feixe paralelo incide no espelho (5) a 45° , passa por uma lente móvel (6), pelo divisor de feixe (8), por uma outra lente (7) e é, então, focalizado na retina do olho.

Figura 3.9 - Diagrama esquemático da configuração experimental utilizado por Liang [Lia94].

Nesse primeiro caminho óptico do *laser* o objetivo é gerar um ponto perfeitamente focalizado na fóvea por meio do ajuste na posição da lente (6). A luz refletida de maneira difusa no fundo do olho retorna por todos os componentes ópticos do olho (humor vítreo, cristalino, humor aquoso, córnea), passa pela lente (7), reflete no divisor de feixe (8) e segue o caminho pelas lentes (9), (10) e (12), passando pelo diafragma (11). A finalidade do diafragma é evitar que o reflexo da luz na córnea chegue ao CCD, devido à projeção (7). Finalmente, a frente de onda (*wavefront*) incide no sensor de HS e é focalizada no CCD. A imagem do CCD é digitalizada, armazenada, depois

processada e, finalmente, analisada no computador (PC). As Figuras 3.10, 3.11 e 3.12 apresentam exemplos de imagens adquiridas por esse sistema. A região que cobre os pontos centrais (malha 7 x 7) é a área sobre a qual é feita, posteriormente, a análise dos erros refrativos (indicada pelo quadrado branco na Figura 3.10).

Figura 3.10 - Imagem padrão correspondente a onda plana marcada com a área útil [Lia94].

A Figura 3.10 mostra uma imagem padrão para o teste de uma onda plana. Utilizando um padrão como referência é possível remover automaticamente todos os erros de fase remanescentes do braço de detecção do sistema, o qual é constituído pelas lentes (7), (9), (10) e (12), e pelo divisor de feixe (8). As Figuras 3.11 e 3.12 apresentam o resultado da aquisição das imagens de dois olhos de pessoas diferentes. As duas imagens do mesmo olho do primeiro paciente correspondem a uma miopia de -1.0 a -1.5 D (Figura 3.11) e as do segundo correspondem a um olho emétrope (Figura 3.12). Cada imagem digitalizada é o resultado da média de quatro amostras individuais extraídas no período de 1 minuto sob as mesmas condições. Devido à reflexão da luz do *laser* vinda da lente (7) antes de penetrar no olho, um padrão de distúrbio aparece ao redor do ponto central das imagens. Esse distúrbio é removido posteriormente por um *software* de pré-processamento de imagens, sendo substituído por um pequeno ponto. Seria possível evitar o problema inclinando a lente (7), ou substituindo-a por uma lente com um pequeno raio de curvatura.

Figura 3.11 – Imagens do olho de um paciente L (L1 e L2, respectivamente) obtidas em dois dias consecutivos [Lia94].

Figura 3.12 - Imagens do olho de um paciente B (B1 e B2, respectivamente), obtidas no mesmo dia com diferença de meia hora entre um e outro [Lia94].

3.4 Análise das imagens baseada no polinômio de Zernike

No sistema descrito anteriormente, os padrões de imagem são analisados por um *software* que, inicialmente, acha o centróide (posição x e y central) da imagem completa e localiza o ponto de foco que se encontra mais perto do mesmo. Em seguida, a área completa da imagem é subdividida em pequenas regiões nas quais existe somente um ponto de foco. A posição do ponto focal em cada sub-abertura é determinada novamente pelo centróide. Para determinar a posição dos pontos focais é utilizado o algoritmo centróide [Cox81], cuja eficiência é comprovado e largamente empregado para o rastreamento de pontos em resolução de *sub-pixel* [Sal81]. Comparando as posições dos pontos focais da imagem referência (Figura 3.10) com as posições correspondentes nas imagens adquiridas dos olhos reais (Figuras 3.11 e 3.12) é possível calcular o deslocamento de cada ponto focal em ambas as direções, x e y [Lia94].

Se associarmos ao plano da imagem HS um sistema de coordenadas cartesianas (x , y) podemos, por meio do processamento digital, calcular o centro geométrico (o centróide) de cada ponto e sua respectiva coordenada. Para um sensor HS com n lenticulas teremos uma imagem de referência com n pontos (x_n , y_n) e imagens de olhos amétropes (com aberrações) também com n pontos (x_a , y_a). Então é calculada a derivada parcial dos pontos centrais (malha 7 x 7) da frente de onda deformada ($W(x,y)$) vinda do olho, em ambas as direções x e y , determinadas pelas equações (3.1a) e (3.1b), em que f é o comprimento focal do conjunto de microlentes.

$$\frac{\partial W_n(x,y)}{\partial x} = \frac{x_n - x_a}{f} \quad (3.1a)$$

$$\frac{\partial W_n(x,y)}{\partial y} = \frac{y_n - y_a}{f} \quad (3.1b)$$

O objetivo é reconstruir a frente de onda $W(x,y)$ a partir das derivadas parciais. Muitas abordagens foram propostas na literatura [Cub79] [Sou80] para tal, sendo que Liang [Lia94] usou os polinômios de Zernike, utilizados desde 1934 para descrever matematicamente as anomalias dos sistemas ópticos. O benefício é que o *wavefront* pode ser apresentado em componentes independentes, cada um representando anomalias específicas do globo ocular [Lia97a]. Cada um desses componentes pode, então, ser analisado para a posterior correção das anomalias [Wav00]. Outro ponto que favorece a escolha dos polinômios de Zernike para avaliar a frente de onda é que eles são permutáveis com a expansão bidimensional de Taylor para o grau quatro [Lia94].

Assume-se que a frente de onda $W(x,y)$ pode ser expressa pela equação (3.2), onde $Z_i(x,y)$ denotam os polinômios de Zernike de quarto grau (Tabela 3.1) e C_i são os coeficientes do modo de Zernike.

$$W(x,y) = \sum_{i=0}^{14} C_i Z_i(x,y) \quad (3.2)$$

A partir da equação (3.2) são obtidas as derivadas parciais de $W(x,y)$:

$$\frac{\partial W(x,y)}{\partial x} = \sum_{i=0}^{14} C_i \frac{\partial Z_i(x,y)}{\partial x} \quad (3.3 a)$$

$$\frac{\partial W(x,y)}{\partial y} = \sum_{i=0}^{14} C_i \frac{\partial Z_i(x,y)}{\partial y} \quad (3.3 \text{ b})$$

Tabela 3.1 - Polinômio de Zernike elevado ao quarto grau [Lia94].

<i>Termo</i>	<i>Representação monomial</i>	<i>Significado</i>
$Z_0(x,y)$	1	Termo constante
$Z_1(x,y)$	x	Tilt na direção x
$Z_2(x,y)$	y	Tilt na direção y
$Z_3(x,y)$	$2xy$	Astigmatismo com eixo a $\pm 45^\circ$
$Z_4(x,y)$	$-1+2y^2+2x^2$	Aberração focal (miopia ou hipermetropia)
$Z_5(x,y)$	y^2-x^2	Astigmatismo com eixo a $\pm 90^\circ$
$Z_6(x,y)$	$3xy^2-x^3$	Coma de 3ª ordem no eixo x Coma de 3ª ordem no eixo y
$Z_7(x,y)$	$-2x+3xy^2+3x^3$	
$Z_8(x,y)$	$-2y+3y^3+3x^2y$	
$Z_9(x,y)$	y^3-3x^2y	
$Z_{10}(x,y)$	$4y^3x-4x^3y$	Aberração esférica de 3ª ordem
$Z_{11}(x,y)$	$-6xy+8y^3x+8x^3y$	
$Z_{12}(x,y)$	$1-6y^2-6x^2+6y^4+12x^2y^2-6x^4$	
$Z_{13}(x,y)$	$-3y^2+3x^2+4y^4-4x^2y^2-4x^4$	
$Z_{14}(x,y)$	$y^4-6x^2y^2+x^4$	

Interpolando, pelo método dos mínimos quadrados, os valores das derivadas parciais em (3.1) pelos valores na equação (3.3), pode-se calcular os coeficientes como:

$$C = [MT].[DER] \quad (3.4)$$

em que a matriz de coeficientes é da forma:

$$C_i = [C_0 C_1 C_2 \dots C_{14}]^T \quad (3.5)$$

[DER] é a matriz coluna de derivadas e [MT] é uma matriz de transformação de dimensão $14 \times 2N^2$, onde $2N^2$ é o total de medidas das derivadas. A matriz [MT] é obtida pela inversa das derivadas parciais das expressões que contém os polinômios de Zernike. Conhecendo as derivadas, pode-se calcular os coeficientes de Zernike, a distribuição da frente de onda $W(x,y)$, e a contribuição individual dos modos de Zernike [Lia94]. Com esses dados obtêm-se informações sobre as correções refrativas

convencionais, como *defocus* (miopia, hipermetropia) e astigmatismo, assumindo que a frente de onda da lente esférico-cilíndrica equivalente possa ser escrita como:

$$W_{lente}(x, y) = 2C_3xy + 2C_4(x^2 + y^2) + C_5(y^2 - x^2) \quad (3.6)$$

A partir dos valores dos coeficientes de Zernike C_3 , C_4 e C_5 as equações (3.7) a (3.10) permitem calcular os valores de defocus (ϕ_D), astigmatismo (ϕ_A) e eixo do astigmatismo (α). O termo D indica o tamanho do diâmetro da pupila.

$$\phi_D = \frac{-16C_4}{D^2} - 0.5\phi_A \quad (3.7)$$

$$\phi_A = \frac{\sqrt{\pm 16(C_3^2 + C_5^2)}}{D^2} \quad (3.8)$$

$$\alpha = 90^\circ + 0.5 \tan^{-1}\left(\frac{-C_3}{C_5}\right) \quad \text{para } -C_3/C_5 < 0 \quad (3.9)$$

$$\alpha = 0.5 \tan^{-1}\left(\frac{-C_3}{C_5}\right) \quad \text{para } C_3/C_5 \leq 0 \quad (3.10)$$

Na Tabela 3.2 são mostrados os resultados práticos das medições das imagens apresentadas nas Figuras 3.11 e 3.12 utilizando os polinômios de Zernike [Lia94]. O valores de C_3 , C_4 e C_5 utilizados para o cálculo dessas medidas são apresentados na Tabela 3.3. O valor do diâmetro da pupila (D) foi 5.4 mm.

Tabela 3.2 - Medição do defocus e atigmatismo dos olhos dos pacientes L e B [Lia94].

	L1	L2	B1	B2
Defocus	- 1.25 D	- 1.21 D	- 0.46 D	-0.36 D
Astigmatismo	- 0.54 D	- 0.59 D	- 0.34 D	- 0. 56 D
Eixo do astigmatismo	58°	54°	38°	30°

Tabela 3.3 – Coeficientes de Zernike para os olhos dos pacientes L e B [Lia94].

	L1	L2	B1	B2
C_3	0.89	1.02	0.60	0.87
C_4	2.78	2.75	1.15	1.17
C_5	0.43	0.33	-0.15	-0.52

Em 1997, Liang & Williams [Lia97a] aprimoraram a configuração do *hardware* de aquisição, e em 2000 uma outra configuração foi desenvolvida por Hamam [Ham00]. Este último propôs um método numérico direto para a reconstrução das curvas de anomalias vindas do conjunto de pontos medidos, sem passar por um ajuste da frente de onda. O método é rápido e não requer o conhecimento preciso da posição do centro da pupila do olho a ser medido. O valor máximo das anomalias pode ser estimado em situações extremas, em que a quantidade de anomalias é suficientemente grande de modo que alguns pontos são deslocados para o interior das áreas onde são esperados. O método sugerido também simplifica a discussão sobre a precisão e as limitações das medidas, e requer modificações no sistema de aquisição convencional. A nova configuração óptica reduz as perdas e ruídos, o que permite usar uma baixa potência de luminosidade da luz durante a exposição dos olhos, bem como detectar um padrão óptico que não requer nenhum processamento de imagem adicional.

4. TÉCNICAS DE EXTRAÇÃO DE CARACTERÍSTICAS

Este capítulo introduz as técnicas utilizadas para realizar a extração de características relevantes das imagens do olho humano após o pré-processamento. O objetivo da extração é reduzir a quantidade de dados de entrada do módulo de análise, pois uma grande quantidade de atributos pode inviabilizar um bom desempenho dos algoritmos de AM na fase de generalização dos resultados. Duas abordagens distintas foram implementadas e testadas. Do método de análise multivariada foi implementada a técnica da análise de componentes principais (*Principal Component Analysis* – PCA), e do método de transformada de *wavelet* foi implementada a transformada de Gabor.

Entre as opções de técnicas de análise multivariada optou-se por investigar inicialmente a PCA devido à grande quantidade de publicações [Gua94] [Moh95] [Mou97] [Oli97] [Fit99] [Ton02] que relatam a sua eficiência para o propósito de redução de imagens sem perda de informações relevantes. A PCA foi aplicada às imagens originais descritas no domínio espacial, de forma que o conjunto de atributos de entrada para o módulo de análise corresponde à descrição de uma imagem após a sua manipulação pelo algoritmo da PCA.

Buscando uma alternativa à PCA capaz de permitir uma maior redução no número de atributos de entrada a serem fornecidos para o módulo de análise, foi utilizada a transformada wavelet de Gabor para extrair atributos relevantes das imagens. A opção por utilizar tentativamente esta técnica deveu-se ao conhecimento de aplicações bem sucedidas da mesma em biometria (reconhecimento de indivíduos pela íris [Dau02]) e na recuperação de imagens médicas por conteúdo [Cas03] – esta última em trabalho desenvolvido por pesquisadores do ICMC/USP. Vale ressaltar que existem muitas outras abordagens possíveis para a extração de atributos das imagens, e o tratamento desta questão merece investigação adicional.

A seguir, a técnica de análise de componentes principais (PCA) é introduzida na Seção 4.1, e na Seção 4.2 é apresentada a transformada *wavelet* de Gabor.

4.1 Análise de componentes principais (PCA)

Análise multivariada é o ramo da estatística que estuda o relacionamento das variáveis dependentes em um determinado conjunto de dados que consiste de n observações, cada qual contendo valores para p variáveis [Joh82]. Apesar de existirem muitas técnicas de análise multivariada, todas têm como objetivo comum a redução e simplificação da complexidade do problema em questão. Na prática, essas técnicas podem reduzir o número de variáveis a serem computadas, ou eliminar variáveis difíceis de serem observadas ou medidas, garantindo, ainda assim, que informação relevante não seja perdida. Do ponto de vista teórico, essas técnicas podem reduzir a dimensionalidade dos dados, ainda que acarretem a perda de algumas informações. Dentre os objetivos da análise multivariada, pode-se citar:

- **Redução de dados:** Com a finalidade de facilitar a interpretação, os dados são reduzidos para a forma mais simples possível, sem sacrificar informações relevantes. Isso pode ser feito, por exemplo, transformando um conjunto de variáveis interdependentes em variáveis independentes, ou reduzindo a dimensionalidade.
- **Agrupamento dos dados:** Grupos de variáveis ou objetos “similares” são criados, baseando-se em medidas de similaridade de características.
- **Investigação de dependência entre variáveis:** É estabelecida a natureza do relacionamento entre as variáveis, isto é, se as variáveis são mutuamente independentes ou se existem uma ou mais variáveis que dependem de outras.
- **Previsão:** Os relacionamentos entre as variáveis podem ser determinados para prever os valores de uma ou mais variáveis de interesse.

A Análise de Componentes Principais (*Principal Component Analysis* – PCA), também conhecida como transformada de *Hotelling* ou transformada de Karhunen-Loève, é uma técnica analítica que transforma um grupo de variáveis correlacionadas em variáveis não correlacionadas. A PCA é um dos métodos estatísticos multivariados

mais simples, em que dadas p variáveis $X_1, X_2, \dots, X_p$, busca-se encontrar combinações lineares destas para produzir índices $Z_1, Z_2, \dots, Z_p$ não correlacionados.

A ausência de correlação é uma propriedade útil, pois significa que os índices estão medindo diferentes “dimensões” dos dados. Esses índices são ordenados de forma que Z_1 representa a maior quantidade de variação, Z_2 representa a segunda maior quantidade de variação, e assim por diante. Isto é,

$$\text{var}(Z_1) \geq \text{var}(Z_2) \geq \dots \geq \text{var}(Z_p)$$

onde $\text{var}(Z_i)$ denota a variância de Z_i no conjunto de dados. Os índices Z_i são chamados de componentes principais.

Ao realizar uma PCA espera-se sempre que as variâncias da maioria dos índices sejam tão pequenas que possam ser desprezadas. Nesse caso, a variação dos conjuntos de dados pode ser descrita adequadamente pelas poucas variáveis Z cujas variâncias são consideradas não desprezíveis. Um certo grau de economia é então atingido, já que a variação nas p variáveis originais X fica representada por um menor número de índices Z .

Deve ser enfatizado que nem sempre a PCA reduz um grande número de variáveis originais a um pequeno número de variáveis transformadas. De fato, se as variáveis originais não são correlacionadas, a análise não produz efeito algum. Os melhores resultados são obtidos quando as variáveis originais são altamente correlacionadas, positivas ou negativamente. A PCA é uma das abordagens mais populares para o problema de reconhecimento de faces, e uma das mais adequadas para a extração de características [Oli97] [Rom97] [Cel99].

Para utilizar os resultados de uma PCA não é necessário saber como foram derivadas as equações para gerar as componentes principais. Todavia, é útil entender a natureza dessas equações. De fato, uma PCA resume-se em encontrar os autovalores da matriz de covariância dos dados. Essa matriz é simétrica e tem a seguinte forma:

$$C = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1p} \\ c_{21} & c_{22} & \cdots & c_{2p} \\ \vdots & \vdots & & \vdots \\ c_{p1} & c_{p2} & \cdots & c_{pp} \end{bmatrix}$$

onde o elemento da diagonal principal C_{ii} é a variância de X_i e C_{ij} ($i \neq j$) é a covariância das variáveis X_i e X_j .

As variáveis $X_1, X_2, \dots, X_p$ devem ser padronizadas para que tenham média igual a 0 e variância igual a 1 antes de proceder com a análise para evitar que uma determinada variável tenha influência excessiva nas componentes principais. Dessa forma, depois de feita a codificação dos dados, a matriz de covariância C terá a forma:

$$C = \begin{bmatrix} 1 & c_{12} & \dots & c_{1p} \\ c_{21} & 1 & \dots & c_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ c_{p1} & c_{p2} & \dots & 1 \end{bmatrix}$$

que coincide com a matriz de correlação dos dados, onde $C_{ij} = C_{ji}$ é a correlação entre X_i e X_j . Em outras palavras, a PCA é realizada sobre a matriz de correlação dos dados. Nesse caso, a soma dos elementos da diagonal principal é igual a p , ou seja, ao número de variáveis originais. Nesse caso, o problema de autovalor, comumente encontrado em álgebra linear, resume-se a resolver a equação: $Ca = \lambda a$. O problema tem soluções não triviais, isto é, $a \neq 0$, somente para valores especiais de λ , os quais são chamados de autovalores da matriz de correlação C . Os vetores de a associados aos autovalores são chamados de autovetores. A matriz de correlação é caracterizada por autovalores reais não negativos.

Sejam $\lambda_1, \lambda_2, \dots, \lambda_p$ os autovalores da matriz de correlação C e $a_1, a_2, \dots, a_p$ os respectivos autovetores associados, pode-se então escrever:

$$Ca_j = \lambda_j a_j \quad j = 1, 2, \dots, p$$

Supondo-se que os correspondentes autovalores sejam arranjados em ordem decrescente, de modo que $\lambda_1 = \lambda_{\max}$ e que os autovetores associados sejam usados para construir uma matriz de dimensão $(p \times p)$ pode-se combinar o conjunto das p equações em uma única equação: $CA = A\Lambda$, onde Λ é uma matriz diagonal definida pelos autovetores da matriz C . A matriz A é ortogonal, ou seja, suas colunas, os autovetores de C , representam vetores linearmente independentes. A partir da análise de componentes principais, alguns resultados importantes são obtidos [Man86]:

- As variâncias das componentes principais são iguais aos autovalores da matriz de correlação C , onde λ_i é o autovalor relativo à i -ésima componente principal, $Z_i = a_{i1} X_1 + a_{i2} X_2 + \dots + a_{ip} X_p$. Em particular, $\text{var}(Z_i) = \lambda_i$.
- Da álgebra linear tem-se que, se os autovalores são distintos, então os autovetores correspondentes são linearmente independentes. Esse resultado permite concluir que uma condição necessária para obter a matriz A é que os autovalores sejam distintos. Portanto, a variância de cada componente principal é única.
- As constantes $a_{i1}, a_{i2}, \dots, a_{ip}$ são os elementos do i -ésimo autovetor correspondente. Uma propriedade importante dos autovalores é que eles somados correspondem à soma dos elementos da diagonal principal de C . Isto é, $\lambda_1 + \lambda_2 + \dots + \lambda_p = C_{11} + C_{22} + \dots + C_{pp}$. Desde que C_{ii} é a variância de X_i e λ_i é a variância de Z_i , a soma das variâncias das p componentes principais é igual à soma das variâncias das variáveis originais.

Do ponto de vista do reconhecimento de padrões estatísticos, o valor prático da PCA é fornecer uma técnica efetiva para redução de dimensionalidade. Pode-se reduzir o número de características necessárias para uma representação efetiva dos dados descartando-se combinações lineares com baixas variâncias e retendo-se somente aqueles termos que têm variâncias significativas. Os passos para a utilização da PCA são [Oli97]:

- Normalizar os dados correspondentes às variáveis $X_1, X_2, \dots, X_p$ para que tenham média igual a 0 e variância igual a 1;
- Calcular a matriz de correlação C ;
- Encontrar os autovalores $\lambda_1, \lambda_2, \dots, \lambda_p$ da matriz C e seus correspondentes autovetores $a_1, a_2, \dots, a_p$. O coeficiente da i -ésima componente principal é dado por a_i , enquanto λ_i é a sua variância;
- Descartar as componentes que acumulem uma pequena proporção da variação dos dados. Por exemplo, se os dados originais tiverem 100 variáveis e as 10 primeiras componentes principais forem responsáveis por 95% do total da variância, as outras 90 componentes podem ser ignoradas.

4.2 Transformada wavelet de Gabor

Wavelets são funções que satisfazem requisitos matemáticos utilizados na representação de dados ou outras funções. A idéia de utilizar funções para representar outras não é nova, e existe desde que Fourier descobriu que senos e co-senos podem ser utilizados para representar outras funções. Porém, na análise de *wavelet*, a escala possui um papel importante no processamento dos dados. A *wavelet* pode ser processada utilizando diferentes escalas ou resoluções. Quando um sinal é observado em uma grande “janela” de tempo, suas características observadas são mais grosseiras, e quando é utilizada uma pequena “janela” de tempo as características desse sinal observado são mais finas. Portanto, é possível observar tanto características mais aparentes quanto as menos perceptíveis [Mal99]. A aplicação das *wavelets* é observada em vários segmentos da ciência, como astronomia, engenharia, música, matemática, processamento de sinais e processamento de imagens. No caso de processamento de imagens, um dos pontos fortes da transformada *wavelet* é que é possível representar a imagem toda com uma quantidade mínima de dados. Esses dados podem ser aproveitados na montagem de um vetor de características da imagem.

A transformada *wavelet* de Gabor (*Gabor Wavelets Transform* – GWT) foi originalmente proposta por Gabor [Gab46] [Cha93] [Man96]. Em seu trabalho pioneiro ele demonstrou que a capacidade de caracterizar um sinal simultaneamente nos domínios temporal (ou espacial) e das frequências é limitada pela “razão de incerteza conjunta”, também conhecida como “princípio de Heisenberg”:

$$(\Delta t) (\Delta \omega) \geq 1/4\pi \quad (4.1)$$

onde, Δt e $\Delta \omega$ representam incertezas nos domínios temporal e das frequências, respectivamente. Gabor determinou uma família de funções que atingem esse limite inferior de “incerteza conjunta” - equação (4.2).

$$f(t) = \exp \left[-\frac{1}{2} \left(\frac{t}{\sigma} \right)^2 + i\omega t \right] \quad (4.2)$$

Essa função $f(t)$ descreve uma onda senoidal com frequência ω modulada por um envelope Gaussiano com duração (ou extensão) σ . O conjunto original de filtros proposto por Gabor foi estendido para o caso bidimensional (2D) por Daugman [Dau95], podendo assim ser aplicado em imagens digitais. Em duas dimensões essa função tem a seguinte descrição:

$$f(x, y, u_o, v_o, \sigma_x, \sigma_y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp\left\{-\frac{1}{2}\left[\left(\frac{x}{\sigma_x}\right)^2 + \left(\frac{y}{\sigma_y}\right)^2\right]\right\} \exp\{2\pi i(u_o x + v_o y)\} \quad (4.3)$$

Essa função complexa pode ser melhor entendida pelas suas componentes, real (f_R) e imaginária (f_I):

$$f_R(x, y, u_o, v_o, \sigma_x, \sigma_y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp\left\{-\frac{1}{2}\left[\left(\frac{x}{\sigma_x}\right)^2 + \left(\frac{y}{\sigma_y}\right)^2\right]\right\} \cos(2\pi(u_o x + v_o y)) \quad (4.4)$$

$$f_I(x, y, u_o, v_o, \sigma_x, \sigma_y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp\left\{-\frac{1}{2}\left[\left(\frac{x}{\sigma_x}\right)^2 + \left(\frac{y}{\sigma_y}\right)^2\right]\right\} \sin(2\pi(u_o x + v_o y)) \quad (4.5)$$

onde u_o e v_o representam a frequência espacial implementada no filtro ao longo das direções x e y , respectivamente e σ_x e σ_y a extensão espacial da função de Gabor.

Um melhor entendimento dos filtros de Gabor pode ser obtido examinando a funções (4.3) e (4.4) no domínio de Fourier. A transformada de Fourier da componente real (equação 4.4) é dada por:

$$F_R(u, v, \sigma_x, \sigma_y, u_o, v_o) = \left[\exp\left\{-\frac{1}{2}[\sigma_x^2(u - u_o)^2 + \sigma_y^2(v - v_o)^2]\right\} \right] + \left[\exp\left\{-\frac{1}{2}[\sigma_x^2(u + u_o)^2 + \sigma_y^2(v + v_o)^2]\right\} \right] \quad (4.6)$$

Na Figura 4.1b é ilustrado o comportamento da componente real da equação (4.3). São apresentadas duas exponenciais Gaussianas centradas na frequência selecionada (u_o , v_o) e $(-u_o, -v_o)$ e com variâncias iguais a $1/\sigma_x^2$ e $1/\sigma_y^2$.

Figura 4.1 – Componente real da função de Gabor no domínio, espacial (a) e das frequências (b) [Dau95].

Pode-se observar nas equações (4.4) e (4.6) que a extensão do filtro varia inversamente nos domínios espacial e das frequências. Esses filtros 2D exibem resultados promissores em aplicações de reconhecimento de textura e objetos [Tuc93]. Os vetores de atributos de textura são formados extraíndo-se medidas estatísticas dos padrões de textura filtrados [Dau01]. Medidas estatísticas extraídas das imagens filtradas deverão ser próximas se elas possuírem texturas semelhantes, e valores bem diferentes caso contrário.

A transformada de Gabor mostra-se adequada para a extração de características, pois minimiza a incerteza das articulações bidimensionais no espaço e frequência [Dau88]. Esses filtros 2D podem ser ajustáveis na orientação e na escala. Dessa forma, detectores de linhas e bordas, e cálculos estatísticos de micro características em uma determinada região são usualmente efetuados para caracterizar a informação da textura [Man96]. A técnica de Gabor têm sido utilizada em várias aplicações de processamento de imagens, por exemplo, na representação de imagens mediante a definição de um conjunto de *wavelets* que fornece uma representação completa das mesmas [Lee96], na recuperação de imagens baseadas na textura [Man96] e na recuperação de imagens por conteúdo (CBIR - *Content-based Image Retrieval*) [Cas03]. Também são relatadas

aplicações no reconhecimento de indivíduos pela íris [Dau02], reconhecimento de impressões digitais [Jai01] e de faces humanas em seqüências de vídeo [Fer01].

Capítulo 5

---

5. TÉCNICAS DE APRENDIZADO DE MÁQUINA PARA ANÁLISE DE IMAGENS

Neste capítulo são introduzidas as técnicas de Aprendizado de Máquina (AM) que foram estudadas e posteriormente utilizadas para analisar os dados vindos do módulo de extração de características. AM é a linha de pesquisa em Inteligência Computacional que estuda métodos capazes de extrair conceitos (conhecimento) a partir de amostras de dados [Mit97]. Pesquisadores se inspiraram nos sistemas biológicos para desenvolver as Redes Neurais Artificiais (RNAs), cujos conceitos são baseados nos mecanismos de aprendizado cerebrais; e os Algoritmos Genéticos (AGs), que têm seus conceitos inspirados no processo de evolução natural e na genética. No campo de aprendizado simbólico, têm-se as árvores de decisão e, inspirados nos processos cognitivos, foram desenvolvidos os sistemas de Raciocínio Baseado em Casos (RBC). As Máquinas de Vetores Suporte (*Support Vector Machine* - SVMs), por sua vez, foram fundamentadas em teorias estatísticas [Lor02].

Uma motivação para investigar técnicas de AM para o reconhecimento e análise das imagens HS no contexto deste projeto foi a busca por um sistema de medição capaz de interpretar o conteúdo das imagens como um todo, ao invés de fazer a medição pela comparação de dados discretos extraídos da imagem com dados extraídos de uma imagem de referência, como é feito atualmente por especialistas em Optometria utilizando técnicas como os polinômios de Zernike [Lia94] (Seção 3.4). Nessas abordagens convencionais, muitas informações contidas nas imagens são ignoradas no processo de análise [Him01] [Val01]. Uma abordagem baseada na análise do conteúdo da imagem permitirá, no futuro, que além dos vícios refrativos sejam detectadas outras enfermidades do olho humano a partir das imagens de HS, como é o caso do “*Tears Film Breakup*” [Him01].

Dentre as várias técnicas conhecidas de AM potencialmente aplicáveis, as SVMs possuem características que justificam a sua utilização no contexto deste projeto [Smo99] [Lor02]. As SMVs têm boa capacidade de generalização; robustez diante de objetos de dimensões elevadas, como imagens; convexidade da função objetivo, isto é, possui apenas um mínimo global; capacidade de lidar com dados ruidosos e uma base teórica bem estabelecida na Matemática e Estatística. As SVMs têm sido muito utilizadas em aplicações recentes, com resultados comparáveis aos obtidos por outros algoritmos de aprendizado, como RNAs, tendo se mostrado superiores em algumas tarefas. É o caso de reconhecimento de caracteres [Lec95] [Dum98], detecção de faces humanas em imagens [Osu98] [Fer99] e reconhecimento de pessoas pelas características da cor e forma do corpo [Nak02].

O capítulo está organizado em 4 seções. Na Seção 5.1 são introduzidas noções teóricas sobre as RNAs. Na Seção 5.2 é discutida a sua aplicação na área de reconhecimento e análise de imagens, e na Seção 5.3 é discutida a aplicação de AGs associados às RNAs. Por fim, na Seção 5.4 é introduzida a fundamentação teórica sobre SVMs.

5.1 Redes neurais artificiais

As Redes Neurais Artificiais (RNAs) são sistemas de AM que empregam como unidade de processamento fundamental um modelo de neurônio artificial, inspirado no funcionamento básico dos neurônios biológicos [Tin99]. O adjetivo “neural” é usado porque muito da inspiração para tais redes veio da neurociência [Her91]. Dentre as várias definições de RNA, uma das mais consagradas é a de Hecht-Nielsen [Hec89] [Bre96]:

“Uma Rede Neural Artificial é uma estrutura de processamento de informação distribuída e paralela, consistindo de unidades de processamento interconectadas por canais de comunicação de sinais unidirecionais. Cada unidade de processamento tem uma única conexão de saída que pode se ramificar em tantas conexões colaterais quantas forem desejadas, cada qual transportando o mesmo sinal – o sinal de saída da unidade de processamento. O sinal de saída pode ser de um tipo matemático qualquer. Todo o processamento é local, isto é, depende apenas dos valores correntes do sinal de

entrada que chega por meio das conexões de entrada e dos valores armazenados localmente na unidade de processamento”.

As RNAs foram inspiradas no pouco que se conhece sobre o cérebro humano. Além de se caracterizar pela robustez e tolerância a falhas no processamento, o cérebro pode ser treinado para se ajustar a um novo ambiente, e tem a capacidade de processar informações nebulosas ou ruidosas. Possui, ainda, alto grau de paralelismo no processamento (alta velocidade) e é pequeno, compacto e de baixo consumo de energia. As principais características que tornam as RNAs interessantes para a solução de problemas são, segundo Bittencourt [Bit96]:

- ❑ A capacidade de “aprender” por meio de exemplos e de generalizar esse aprendizado de maneira a reconhecer instâncias similares que nunca haviam sido apresentadas como exemplo;
- ❑ O bom desempenho em tarefas mal definidas, em que falta conhecimento explícito sobre como encontrar uma solução;
- ❑ Não requerem conhecimento de eventuais modelos matemáticos dos domínios de aplicação;
- ❑ A elevada imunidade ao ruído, isto é, apesar de piorar gradativamente, o desempenho das redes não entra em colapso na presença de informações falsas ou ausentes, como é o caso de programas convencionais.

5.1.1 Fundamentos

Uma RNA consiste de um conjunto de neurônios (células) conectadas por uma regra de propagação. Cada neurônio recebe suas entradas com os pesos associados, vindos de outros neurônios, ou de um estímulo externo. Para cada neurônio da rede, uma função de ativação é aplicada sobre a somatória ponderada de suas entradas. A camada de entrada possui um neurônio especial, denominado *bias*, que serve para aumentar os graus de liberdade, permitindo uma melhor adaptação da rede ao conhecimento a ela fornecido [Tat00]. O estado de ativação de um neurônio é determinado pela função de ativação que, geralmente, é uma função sigmoideal ou uma função degrau (*Hard Limiter*) [Per94]. A função degrau requer menor tempo computacional, porém pode introduzir erros. A função de saída é geralmente uma

função identidade. Uma regra de propagação consiste das funções de ativação e de saída. Essa saída trafega pela sinapse até o próximo neurônio. Grande parte das RNAs têm neurônios organizados em camadas, cada qual agrupando neurônios com funções ou propriedades semelhantes. As características das arquiteturas incluem os tipos de conexões, os esquemas de conexão, e as configurações das camadas.

Há basicamente dois tipos de conexões: as excitatórias, que incrementam a atividade de um neurônio e são representadas por sinais positivos, e as inibitórias, que decrementam uma atividade e são representadas por sinais negativos. Tais conexões não precisam ser mutuamente exclusivas; um neurônio pode receber conexões com pesos negativos e outras com pesos positivos.

Os três esquemas básicos de interconexão são: intracamada ou colaterais (conexões entre neurônios de uma mesma camada), intercamada (conexões entre neurônios de diferentes camadas) e recorrentes (conexões com origem e destino no mesmo neurônio). Os sinais de conexões intercamadas se propagam por alimentação progressiva (para frente) ou regressiva (para trás). Sinais de alimentação progressivos permitem que a informação flua em apenas uma direção, os regressivos possibilitam à informação fluir em uma direção e/ou recursivamente [Bre96].

As configurações das camadas e a alimentação regressiva e progressiva combinam neurônios, fluxo de informação e esquemas de conexão em uma arquitetura coerente. Uma camada que recebe sinais externos do ambiente é denominada camada de entrada; e uma camada que emite sinais para o ambiente é denominada camada de saída. Quaisquer outras são denominadas camadas ocultas, e não se conectam diretamente com o ambiente [Car99]. Na rede chamada *feedforward* a propagação do sinal intercamadas é progressiva, e toda entrada é recebida da camada adjacente, com exceção da primeira, em que os sinais de entrada são as próprias entradas da rede. Na fase de treinamento há a determinação e a correção dos pesos e dos *bias* [Bra98].

O objetivo da aprendizagem, ou treinamento, é produzir um conjunto de saídas desejadas ou, no mínimo consistentes, a partir de um conjunto de entradas. Conjuntos de entrada ou saída são referidos como vetores, e a aprendizagem se dá por meio da aplicação sequencial de vetores de entrada e o ajuste do estado interno da rede de acordo com o algoritmo de aprendizagem. Os paradigmas de aprendizado são classificados em três categorias: aprendizagem supervisionada, aprendizagem não supervisionada e de reforço [Bre96], discutidas a seguir.

- **Aprendizado Supervisionado:** Informação sobre a classe a que pertence cada um dos padrões de entrada é conhecida durante o treinamento. A aprendizagem supervisionada é um processo que incorpora um instrutor (ou professor) externo ou uma informação global. São fornecidos à rede pares de treinamento, representando os vetores de entrada e as saídas desejadas. Para cada vetor de entrada aplicado é calculada a saída e comparada ao vetor de saída desejado. O erro é propagado pela rede e os pesos são alterados por um algoritmo que busca minimizar este erro. Técnicas de aprendizagem supervisionada devem decidir quando encerrar o processo; como e por quanto tempo apresentar os pares de treinamento; e suprir informação de erro e desempenho.
- **Aprendizado por reforço:** É um caso particular de aprendizado supervisionado em que a rede é recompensada em caso de acerto e punida em caso de fracasso, não sendo fornecidas à rede as saídas desejadas. A principal diferença em relação ao treinamento supervisionado clássico é a medida de desempenho adotada. No treinamento por reforço a única informação de realimentação fornecida é se uma determinada saída está correta ou não, isto é, não é fornecida a resposta correta para o padrão de entrada. Esse treinamento é uma forma de aprendizado *on-line* obtido por mapeamento de entrada-saída por meio de um processo de triagem e erro desenvolvido para maximizar o índice de desempenho escalar, chamado sinal de reforço.
- **Aprendizado Não Supervisionado:** O aprendizado acontece sem a indicação da classe a qual pertence cada padrão de entrada, e a rede deve agrupar os padrões fornecidos segundo um critério de similaridade definido. A partir do momento em que identifica as regularidades estatísticas do vetor de entrada, a rede desenvolve uma habilidade de formar representações internas para codificar características da entrada e criar novas classes ou grupos automaticamente [Bra98]. A aprendizagem não supervisionada, também conhecida como auto-organização (*self-organization*) não incorpora o instrutor externo e se baseia apenas em informação local e controle interno. O algoritmo de treinamento modifica a rede de forma a produzir vetores de saída consistentes, ou seja, tanto a aplicação de um dos vetores de treinamento quanto de um vetor suficientemente similar irão produzir o mesmo padrão de saída.

O algoritmo de treinamento para aprendizagem supervisionada mais conhecido é o *Backpropagation*. Segundo Lippmann [Lip87], esse algoritmo permite um ajuste de pesos em cada uma das camadas da rede e é projetado para minimizar a soma do quadrado do erro (SSE) (ou o erro médio quadrático – MSE) entre a saída calculada por uma arquitetura multicamadas e a saída desejada. O componente essencial do algoritmo é o método iterativo que propaga regressivamente o erro requerido para adaptar os pesos, a partir dos neurônios da camada de saída para os nós das camadas precedentes. O processo de treinamento começa com a definição de um conjunto arbitrário de pesos para as conexões da rede e envolve duas fases distintas. Na primeira, um vetor de treinamento com a saída desejada é apresentado à rede e propagado para computar uma saída para cada elemento de processamento. As saídas dos nós da última camada são, então, comparadas às saídas desejadas e, a partir disso, calculados os termos de erro. A segunda fase envolve um retrocesso pela rede, a partir da última camada, durante o qual o erro é repassado para cada elemento de processamento e os pesos correspondentes são convenientemente alterados. Em um treinamento bem sucedido, o erro diminui com o aumento do número de iterações e o procedimento converge para um conjunto estável de pesos [Gal00].

Alguns termos freqüentes nesse contexto são: *overtraining* – quando a rede se especializa no conjunto de dados do treinamento e perde a capacidade de generalização (degeneração da generalização); *underfitting* – a rede não converge durante o treinamento, e *overfitting* – a rede memoriza os padrões de treinamento, gravando suas peculiaridades e ruídos, ao invés de extrair as características gerais que permitirão a generalização ou reconhecimento de padrões não vistos durante o treinamento. Alternativas para evitar o *overfitting* são encerrar o treinamento mais cedo, quando o erro de validação começar a subir, ou reduzir, ou podar, os pesos (*prunning*).

Na próxima seção são apresentados os modelos de RNAs *Multilayer Perceptron* (MLP) e *Radial Basis Function* (RBF).

5.1.2 Multilayer perceptron (MLP) e radial basis function (RBF)

O *MultiLayer Perceptron* (MLP) [Rum86] é o modelo de RNA mais popular. Como o nome indica, os neurônios são distribuídos em pelo menos uma camada

intermediária e uma camada de saída. Nas redes com múltiplas camadas, cada camada tem uma função específica. A camada de saída recebe os estímulos da última camada intermediária e constrói o padrão que será a resposta da rede. As camadas intermediárias funcionam como extratores de características, seus pesos codificam características apresentadas nos padrões de entrada e permitem que a rede crie sua própria representação do problema. Uma camada intermediária é suficiente para aproximar qualquer função contínua ou booleana; duas camadas são suficientes para aproximar qualquer função matemática; e três camadas ou mais podem facilitar o treinamento. Não existe uma regra para especificar o número de neurônios nas camadas intermediárias, a escolha depende do número de exemplos de treinamento, da quantidade de ruído presente nos exemplos, da complexidade da função a ser aprendida, e da distribuição estatística dos dados [Car99]. O *Backpropagation* é o algoritmo comumente utilizado para treinar redes MLP. Outros algoritmos de treinamento usados são o *RProp* (*Resilient Backpropagation*) [Rie93] e o *Quickprop* [Fah88]. O *Rprop* é um esquema de aprendizado adaptativo local, que executa um aprendizado supervisionado em *batch*. O *Quickprop* acelera o aprendizado do *Back-propagation* utilizando informações sobre a curvatura da superfície de erro.

A arquitetura das redes do tipo RBF (*Radial Basis Function*) foi proposta por Broomhead & Lowe em 1988 [Bro88]. Geralmente, elas são utilizadas em problemas de aproximação de funções e reconhecimento de padrões. Normalmente, esse tipo de rede possui uma única camada intermediária cujos neurônios utilizam funções de base radial, sendo a função Gaussiana a mais comum [Bra00]. Redes RBF são treinadas por métodos híbridos divididos em dois estágios, um não supervisionado e outro supervisionado. O primeiro estágio determina o número de funções radiais e seus parâmetros por métodos não supervisionados. O segundo ajusta os pesos dos neurônios de saída por meio de modelos lineares como a regra delta [Bra00]. O algoritmo mais utilizado no primeiro estágio de treinamento não supervisionado é o agrupamento *k*-médias. Esse algoritmo agrupa o conjunto de entradas em subconjuntos disjuntos e posiciona os centros no ponto médio de cada um desses subconjuntos, com a finalidade de minimizar a soma dos quadrados das distâncias entre os vetores de entrada e o centro correspondente. Já o segundo estágio de treinamento utiliza um algoritmo para ajustar os pesos nas camadas de saída, uma vez que a transformação do espaço das unidades intermediárias para o espaço de saída é linear [Fac01].

Uma importante característica das redes RBF é observada utilizando-as para classificar padrões. Suponha um conjunto de dados dividido em três classes. A rede MLP pode separá-lo gerando hiperplanos que dividem as classes, já a rede RBF separa as classes em núcleos, isto é, em regiões, utilizando hiper-elipsóides para dividir o espaço de padrões. A divisão do espaço de padrões em hiper-elipsóides possibilita maior precisão na classificação dos padrões de entrada. A seguir são apresentadas semelhanças e diferenças entre os modelos clássicos de redes RBF e MLP. A discussão refere-se, especificamente, às redes RBF com função de ativação Gaussiana usando treinamento híbrido, e às redes MLP treinadas com o *backpropagation* [Lac98].

- A principal diferença entre os dois modelos está no cálculo do valor de ativação da unidade intermediária. Na rede MLP, tal ativação é função do produto escalar entre o vetor de entrada e o vetor peso das conexões que chegam à unidade intermediária. Na rede RBF é uma função da distância euclidiana entre o mesmo vetor de entrada e o vetor peso das conexões que chegam à unidade intermediária da rede.
- Na rede RBF a camada intermediária realiza um mapeamento não linear do espaço de entrada para o espaço da camada intermediária, enquanto a camada de saída é linear, isto é, realiza um mapeamento linear do espaço da camada intermediária para o espaço de saída. Na rede MLP ambas as camadas, intermediária e de saída, realizam mapeamentos não lineares.
- Na rede RBF o valor de ativação da camada intermediária é constante sobre hiper-elipsóides no espaço de entrada. Na rede MLP o valor de ativação da camada intermediária é constante sobre hiperplanos no espaço de entrada.
- A rede RBF possui representação local do vetor de entrada no espaço das unidades intermediárias, ou seja, apenas algumas unidades intermediárias contribuem para o valor de saída. A rede MLP possui representação distribuída do vetor de entrada no espaço das unidades intermediárias, ou seja, todas as unidades intermediárias contribuem para o valor de saída.
- Na rede RBF os parâmetros da rede são determinados em dois estágios de treinamento, um não supervisionado e um supervisionado. Na rede MLP, os parâmetros da rede são determinados simultaneamente em um único treinamento supervisionado. A divisão do treinamento em dois estágios rápidos na RBF evita a lenta repropagação do erro que ocorre no treinamento das redes MLP.

- Justificativas teóricas garantem que redes MLP com número suficiente de unidades e parâmetros treinados apropriadamente podem aproximar qualquer função contínua com precisão arbitrária [Hay94]. As redes RBF também possuem essa propriedade, denominada de aproximação universal de função. Assim, teoricamente sempre existe uma rede RBF equivalente a uma rede MLP, e vice-versa.
- As redes RBF constróem aproximadores locais de funções, e não apresentam resposta significativa em regiões onde não há padrões de treinamento. As redes MLP constróem aproximadores globais de funções, sendo capazes de generalizar em regiões onde não há padrões de treinamento.

5.2 RNAs para reconhecimento e análise de imagens

Técnicas estatísticas tradicionais para reconhecimento de padrões, como discriminante Bayesiano e janelas de Parzen, foram populares até o começo da década de 1990 [Vas95]. Desde então, as RNAs tornaram-se uma alternativa às técnicas clássicas de classificação e agrupamento de padrões. As RNAs são adequadas para aplicações em reconhecimento de formas, síntese de fonemas, filtragem de sinais ruidosos sem atraso de tempo, percepção de objetos escondidos em cenas complicadas, ou então no reconhecimento de relações sutis que não são óbvias para observadores humanos. Além disso, elas são capazes de identificar relações entre dados que são difíceis de se descrever ou desconhecidas até então [Per95]. RNAs podem generalizar, ou seja, podem extrair uma saída útil de um conjunto de dados imperfeito, incompleto ou com ruído, uma característica extremamente útil em sistemas com relação sinal/ruído muito deteriorado. Elas podem capturar funções matemáticas não-lineares e de ordem elevada entre as variáveis de entrada. Sua tolerância ao ruído é um atributo de seu efeito de medianização e processamento paralelo dos dados de entrada [Leo94] [Sim98]. Lee et al. [Lee90] relatam que as RNAs desempenham a tarefa de classificação de imagens tão bem, ou melhor, do que as técnicas estatísticas, uma vez que não requerem que a natureza paramétrica dos dados a serem classificados seja explicitada.

Existem, atualmente, vários modelos de RNAs desenvolvidos para as mais diversas aplicações em processamento e análise de imagens. Esses modelos estão sendo utilizados para executar desde tarefas de processamento de baixo nível, como supressão e ampliação, até para identificar soluções satisfatórias em problemas complexos de

otimização [Egm02]; passando pelo reconhecimento de áreas de superposição [Lip87]; segmentação baseada em características e reconhecimento de objetos [Pal93] [Lus93]. As redes do tipo *self-organising feature map* (SOM) são uma alternativa interessante às técnicas supervisionadas quando não se possui um conjunto de imagens suficientes para gerar um padrão, pois podem aprender a discriminar, por exemplo, diferentes texturas, se fornecido um grupo adequado de características.

Em 2002 um grupo de pesquisadores publicou um levantamento de mais de 200 trabalhos descrevendo aplicações de RNAs no reconhecimento e análise de imagens digitais [Egm02]. As diferentes aplicações foram organizadas em um arcabouço chamado cadeia de processamento de imagem (*image processing chain*). As tarefas que integram essa cadeia são apresentadas na Figura 5.1 e, posteriormente, explicadas.

Figura 5.1 – Tarefas de uma cadeia de processamento de imagens [Egm02].

- **Pré-processamento/filtragem:** Operações que geram como resultado uma imagem modificada que mantém as dimensões da imagem original, por exemplo, uma filtragem para redução de ruído.
- **Redução de dados/Extração de características:** Operações que extraem componentes significativos de uma imagem ou de uma janela da mesma. O número de características extraído é, geralmente, menor do que o número de *pixels* de uma janela de entrada.
- **Segmentação:** Operação que extrai de uma imagem regiões coerentes em relação a algum critério pré-estabelecido. Um exemplo é a separação de diferentes texturas em regiões.

- **Deteção e reconhecimento de objetos:** Determinação de uma posição, e possivelmente, de uma orientação e escala de objetos existentes na imagem e classificação desses objetos.
- **Entendimento/Interpretação:** Obtenção de conhecimento de alto nível (semântico) embutido em uma imagem.
- **Otimização:** Minimização de uma função critério a qual pode ser usada, por exemplo, para *graph matching* ou delineação de objetos (*object delineation*). As técnicas de otimização não constituem propriamente um passo da cadeia de processamento de imagem, mas um conjunto de técnicas auxiliares que dão suporte aos outros passos.

O desempenho das tarefas de uma cadeia de processamento de imagens e sua capacidade de processamento é, de certo modo, determinado pelo nível de abstração dos dados de entrada. Na Tabela 5.1 podem ser observados os diferentes níveis de abstração possíveis. A Tabela 5.2 exibe um panorama da distribuição quantitativa dos 235 trabalhos revisados por Egmonet-Petersen et al. [Egm02] segundo a tarefa realizada pelas RNAs na cadeia de processamento e o nível de abstração dos dados de entrada. Na Tabela 5.3 são identificados os diferentes tipos de RNAs aplicadas às diferentes tarefas da cadeia de processamento de imagens.

Tabela 5.1 – Níveis de abstração dos dados de entrada para a cadeia de processamento de imagens [Egm02].

Nível de abstração	Definição
<i>Pixel</i>	Entradas são as intensidades individuais dos <i>pixels</i> ;
Característica local	Entrada é um conjunto de características derivadas dos <i>pixels</i> ;
Estrutura (borda)	Entrada é a localização relativa de uma ou mais características (<i>edge</i> , <i>corners</i> , junções, superfícies, etc);
Objeto	Entrada são as propriedades individuais do objeto;
Conjunto de objetos	Entrada é a ordem mútua e localização relativa dos objetos detectados;
Caracterização do cenário	Entrada é uma descrição completa da imagem incluindo condições de iluminação, contexto, etc.

Tabela 5.2 - Número de aplicações em que as RNAs realizam uma tarefa específica da cadeia de processamento de imagens usando um determinado nível de abstração [Egm02].

Nível de abstração	Pré-processamento (Reconstrução, Restauração e Ampliação)	Compressão/ Extração de características – (C.C.)	Segmentação	Reconhecimento	Entendimento da imagem	Otimização
<i>Pixel</i>	26	25	39	51	3	5
Características locais	4	2	19	38	2	3
Estrutura (borda)			2	6		5
Objeto						1
Conjunto de objeto				2	2	
Caracterização do cenário				2	2	
Total	30	27	60	97	7	14

Tabela 5.3 - Diferentes tipos de redes neurais aplicadas às várias tarefas na cadeia de processamento de imagens [Egm02].

Tipos de RNA	Tarefas de reconhecimento de padrões									
	Pré- Processamento			C. C.		Segmentação		Reconhecimento		
	Reconstrução	Restauração	Ampliação (<i>enhancement</i>)	Compressão	Extração de características	Segmentação baseada em <i>Pixel</i>	Segmentação baseada em características	Reconhecimento de objetos baseado em <i>pixel</i>	Reconhecimento de objeto baseado em características	Entendimento
<i>Feed-forward</i> - regressão	•	•	•	•	•		•			•
<i>Feed-forward</i> – auto associação				•	•			•		
<i>Feed-forward</i> – classificação			•			•	•	•	•	•
<i>Feed-forward</i> – divisão de pesos								•		
<i>Feed-forward</i> – recursivo							•	•		
<i>Perceptron</i>					•					•
<i>Radial Basis Function</i> (RBF)				•		•	•			
<i>Self-organising feature map</i> (SOM)				•	•	•	•	•	•	•
<i>Learning vector Quantization</i>				•			•			
<i>Holpfield</i>	•	•	•		•	•	•	•	•	•
<i>Cellular</i> (CNN)		•				•	•			
<i>Generalized adaptive neural filters</i>		•								
<i>Adaptive resonance theory</i> (ART)			•					•		
<i>Associative memories</i> (RAM)						•		•	•	•

ADALINE	•										
Neocognitron								•			
Probabilistic						•					
Neural decision tree										•	
Neural belief network										•	
Higher order network								•			
Counterpropagation network							•				
Fuzzy neural/ Neuro-fuzzy system		•			•	•		•	•		
Outros		•	•	•	•	•	•	•	•		•

Este trabalho se enquadra como uma tarefa de Entendimento/Interpretação da Imagem (*Image understanding*), uma área bastante complexa que, normalmente, envolve a junção de técnicas de segmentação ou reconhecimento de objetos com o conhecimento do conteúdo da imagem [Egm02]. A meta do trabalho foi aplicar técnicas de AM para interpretar as imagens de entrada e gerar uma medida dos erros refrativos presentes. Em termos do nível de abstração dos dados de entrada é fornecido um vetor de características extraído da imagem original.

Um trabalho citado na revisão que segue essa linha utiliza RNAs para a análise automática de imagens de cromossomos humanos. Redes do tipo MLP em combinação com conhecimento do domínio (*background knowledge*) são utilizadas para classificar os cromossomos a partir da localização relativa de determinadas características dos mesmos, como borda, junções, etc. [Ler98]. Em outras aplicações, RNAs são utilizadas na classificação de navios para auxiliar o controle de tráfego naval, a partir de imagens obtidas por radares utilizando como nível de abstração de entrada um vetor de características [Pas98]; e para analisar imagens de uma câmera para controle de posicionamento de um robô a partir de características locais dos *pixels* [Wel96]. Reinus et al. [Rei94] e Stassopoulou et al. [Sta96] apresentam trabalhos genuinamente de interpretação de imagens. Reinus treina uma RNA para diagnosticar tumores no osso, e Stassopoulou et al. analisam regiões em imagens de satélite, detectando objetos como pedra, áreas com erosão, etc. para prever o grau de desertificação de uma região. O nível de abstração é um conjunto de objetos, e eles utilizaram *neural decision trees* [For98], modelo semântico baseado em extração de estruturas (borda, junções, superfícies, etc.), e *neural belief networks* [Sta96], que são utilizadas para representar conhecimento do conteúdo de imagens. Esse conhecimento pode ser utilizado para

restringir o número de possíveis interpretações de objetos simples, assim como para reconhecer diferentes configurações de objetos.

O principal problema na aplicação de RNAs em tarefas de Entendimento/ Interpretação é sua característica de “caixa-preta”. É virtualmente impossível explicar porque uma interpretação particular da imagem é melhor do que outra ou justificar os resultados obtidos. Outro problema está relacionado aos dados de entrada. Quando poucos dados podem ser fornecidos contendo características relacionadas à estrutura (bordas, arestas, cantos, junções, superfícies, etc.) ou posicionamento dos objetos, um grande número de imagens é necessário para estabelecer estatisticamente um conjunto adequado de treinamento e testes [Egm02].

5.3 Algoritmos genéticos associados às RNAs

Os Algoritmos Genéticos (AGs) são métodos de busca e otimização introduzidos por John Holland da década de 70 [Hol75]. Tais algoritmos foram inspirados nos mecanismos de evolução de populações de seres vivos, baseados no princípio da seleção natural e sobrevivência do mais adaptado introduzido em 1859 pelo naturalista e fisiologista inglês Charles Darwin em seu livro “A origem das espécies”. Sinteticamente, a teoria da evolução diz que o meio ambiente seleciona, em cada geração, os seres vivos mais aptos de uma população. Durante a reprodução ocorrem fenômenos como mutação e *crossover* (recombinação), entre outros, que atuam sobre o material genético armazenado nos cromossomos. Esses fenômenos mantêm grande variabilidade de seres vivos na população. Sobre essa população diversificada age a seleção natural, permitindo a sobrevivência somente dos seres mais adaptados.

Um AG processa uma população de cadeias de *bits* (ou alguma outra estrutura de dados) que codificam em notação binária o conjunto de parâmetros do problema a ser otimizado. O conjunto de todas as configurações que a cadeia de *bits* pode assumir é denominado espaço de busca. Por exemplo, se a cadeia de *bits* codifica n parâmetros a serem otimizados, então o espaço de busca é um espaço n -dimensional. Tipicamente, um AG inicia com uma população $S(0)$ constituída de N cadeias de *bits* aleatórias (cromossomo). Em seguida, cada cadeia é avaliada. O resultado da avaliação é um valor denominado aptidão que mede a utilidade da solução representada na cadeia de *bits* para resolver um dado problema. As melhores cadeias de *bits* de $S(0)$ são selecionadas e

modificadas por operadores genéticos como *crossover* e mutação. A nova geração de N cadeias de *bits* modificadas constituirá a nova população $S(1)$. O processo é repetido, e uma seqüência de populações $S(2)$, $S(3)$, ..., $S(t-1)$, $S(t)$ é gerada.

A combinação do processo de seleção com os operadores *crossover* e mutação guia o algoritmo ao encontro das melhores cadeias de *bits* (cromossomos), ao mesmo tempo em que explora novas regiões do espaço de busca. Finalmente, o algoritmo converge para uma população $S(t)$ que contém cadeias de *bits* que representam o máximo global (solução ótima) do problema. O *crossover* é aplicado com uma dada probabilidade, denominada de “taxa de *crossover*” a cada par de cromossomos, gerando dois novos cromossomos. O operador de mutação é, geralmente, aplicado após o *crossover*, em cada *bit* dos novos cromossomos gerados. A mutação inverte os valores dos *bits*. Cada bit tem uma probabilidade pequena de ser invertido, por exemplo, taxa de mutação igual 0,001. AGs têm sido aplicados com sucesso em variadas áreas, como processamento de imagens, *layout* de circuitos integrados, otimização combinatorial, cálculo de estruturas de pontes; aprendizado de máquina, otimização de funções numéricas não diferenciáveis, ruidosas e multimodais, e associados a RNAs [Lac01].

Existem várias maneiras de combinar AGs e RNAs: na otimização da topologia de uma rede para otimizar o número de unidades, o número de camadas e as conexões entre unidades; no treinamento da rede, substituindo a regra de aprendizado; e na otimização da regra de aprendizado, ou seja, dada uma topologia, encontrar uma regra de aprendizado eficiente [Lac98]. As classificações encontradas na literatura identificam duas principais abordagens de combinação de AGs com RNAs, de acordo a representação do cromossomo: codificação direta, forte ou de baixo nível; e codificação indireta, fraca ou de alto nível [Roo97]. As técnicas de codificação direta codificam diretamente os parâmetros da rede neural (conexões, pesos, número de unidades, etc.) no cromossomo. A codificação indireta, ao contrário, codifica regras ou estruturas que contém informações ou “receitas” de como construir a rede.

Neste projeto foram utilizados AGs para otimizar a topologia das redes neurais (codificação direta) em vista de sua adequação para tratar alguns dos problemas encontrados na busca por topologias adequadas:

- Como não há limite para o número de unidades e conexões, não é possível uma busca exaustiva por todas as possíveis topologias em um tempo computacional aceitável;

- Existem vários pontos de máximo, ou seja, duas topologias completamente distintas podem ter bom desempenho;
- Pequenas mudanças na topologia podem alterar significativamente o desempenho;
- Em redes treinadas com o *backpropagation* a função objetivo é ruidosa porque a inicialização dos pesos é aleatória.

5.4 Máquinas de vetores suporte

As Máquinas de Vetores Suporte ou SVMs (*Support Vector Machines*) constituem uma família de algoritmos de AM que vem recebendo grande atenção nos últimos anos [Sch98]. Elas foram desenvolvidas pela equipe do pesquisador Vladimir Vapnik [Vap92], e seus fundamentos são provenientes da teoria de aprendizagem estatística apresentada nas décadas de 60 e 70 por Vapnik e Chernovenkis [Vap68]. Vapnik idealizou o princípio indutivo de Minimização do Risco Estrutural (*Structural Risk Minimization* – SRM) que busca minimizar o erro do conjunto de treinamento (risco empírico), juntamente com o erro do conjunto de teste. Segundo Vapnik, a motivação para esse princípio surgiu da necessidade de desenvolver limites teóricos para a capacidade de generalização dos sistemas de aprendizagem. Uma maior generalização normalmente implica em um maior número de acertos na fase de teste. No caso do princípio de Minimização do Risco Empírico utilizado em RNAs a generalização é obtida pela escolha de pesos para os neurônios de modo a prover um erro de treinamento mínimo [Vap95]. No entanto, somente a Minimização do Risco Empírico não fornece bons resultados. É preciso considerar o Risco Estrutural, pois quanto maior esse risco, mais complexo é o hiperplano de decisão dessas funções no espaço de entrada dos dados [Che98]. Obter um equilíbrio entre ambos os erros permite superar tendências de excesso de ajustes (*overfitting*) e melhorar a capacidade de generalização [Car02].

As SVMs podem ser consideradas máquinas de aprendizagem com apenas uma camada escondida, treinadas por um algoritmo supervisionado. Seu treinamento é relativamente simples, e requer a resolução de um problema quadrático, dependente dos vetores de treinamento e de alguns parâmetros especificados pelo usuário. É feito um mapeamento não-linear dos dados em um espaço de dimensão elevada onde um hiperplano ótimo (com a maior margem de separação possível) possa ser construído,

permitindo a separação das classes positivas e negativas ou a utilização da máquina para regressão. Características importantes das SVMs que justificam sua utilização são [Smo99] [Lor02]:

- **Boa capacidade de generalização:** classificadores gerados por uma SVM, em geral, alcançam bons resultados em termos de generalização. Essa capacidade é medida por sua eficiência na classificação de dados não pertencentes ao conjunto de treinamento, portanto, evita o *overfitting*.
- **Robustez em grandes dimensões:** Uma SVM é robusta diante de objetos de grandes dimensões, por exemplo, imagens. É comum a ocorrência de *overfitting* em classificadores para esses tipos de dados gerados por outros métodos inteligentes.
- **Convexidade da função objetivo:** A aplicação de uma SVM implica na otimização de uma função quadrática, que possui apenas um mínimo global. Essa é uma vantagem sobre RNAs, cuja função objetivo a ser minimizada, possui mínimos locais.
- **Teoria bem definida:** As SVMs possuem uma base teórica bem estabelecida na Matemática e Estatística.

Apesar da teoria ser relativamente antiga, as principais aplicações práticas datam da década de 90. Desde 1995 as SVMs têm recebido a atenção da comunidade de reconhecimento de padrões, e oferecem oportunidades de pesquisa interessantes para teóricos e não teóricos. Novas abordagens para a teoria de aprendizagem estatística ainda estão em desenvolvimento, e métodos eficientes para a programação de SVMs estão em estudo. Elas têm sido aplicadas na solução de diversos tipos de problemas, com resultados comparáveis aos obtidos por outros algoritmos de aprendizado, como RNAs, tendo se mostrado superiores em algumas tarefas. É o caso de reconhecimento de caracteres [Lec95] [Dum98], detecção de faces humanas em imagens [Osu98] [Fer99] e reconhecimento de pessoas pelas características da cor e forma do corpo [Nak02]. Scholkopf et al. apresentam uma comparação entre SVMs e redes RBF [Sch96], e mais informações podem ser obtidas de outras fontes [Sau98] [Fac01] [Ker02] [Rup02] [Col02a] [Col02b]. Recentemente, foi organizado em 2002, em paralelo à 16ª *Conferência Internacional de Reconhecimento de Padrões*, o *Workshop Internacional sobre Reconhecimento de Padrões Utilizando SVMs* (SVM2002) [Lee02]. O objetivo do encontro foi discutir o impacto prático e a relevância das SVMs na área

de reconhecimento de padrões. O evento recebeu 57 artigos de 21 países, tendo sido aceitos 30 artigos, sendo um do Brasil [San02]. Dentre os artigos publicados destaca-se uma revisão bibliográfica sobre aplicações de SVMs em reconhecimento de padrões [Byu02].

Apesar do bom desempenho em diversas aplicações, algumas ressalvas devem ser feitas com relação à complexidade computacional na busca por soluções, da ordem de n^2 a n^3 , em que n é o número de exemplos de treinamento. Apesar de polinomial, esse valor pode ser significativo em aplicações em que o volume de dados é grande. Porém, os tempos de treinamento são consideravelmente menores do que em outros algoritmos de aprendizagem. Outro problema é a forma como o conhecimento é adquirido e codificado pelos classificadores gerados por SVMs. Esse conhecimento não é facilmente interpretado, o que é uma desvantagem sobre outras técnicas de AM como as árvores de decisão, que organizam o conhecimento adquirido em uma estrutura diretamente interpretável [Cri01].

Capítulo 6

---

6. PROCESSAMENTO E EXTRAÇÃO DE CARACTERÍSTICAS

Neste capítulo são apresentados o banco de imagens de Hartmann-Shack (HS) utilizado e a implementação dos módulos de pré-processamento de imagens (Etapa 1) e extração de características (Etapa 2) do sistema desenvolvido. O processo de criação do banco de imagens oftalmológicas utilizado na realização de experimentos com diferentes técnicas de AM para o módulo de análise é relatado na Seção 6.1. Em seguida, na Seção 6.2, é descrita a concepção e implementação do módulo de pré-processamento de imagens, cujo objetivo é identificar e extrair as regiões de interesse de forma a padronizar as imagens e, posteriormente, eliminar ruídos que poderiam distorcer os resultados do processo de extração de características. Foram implementados algoritmos inspirados no método de análise de histograma e que utilizam informações espaciais e geométricas do domínio da aplicação. Por fim, na Seção 6.3 é descrita a implementação do módulo de extração de características utilizando o algoritmo da PCA [Jon82] e, posteriormente, a transformada *wavelet* de Gabor [Gab46], cujo objetivo é diminuir a quantidade de atributos de entrada para o módulo de análise.

6.1 Banco de imagens

As imagens foram adquiridas pelo medidor de anomalias baseado nos princípios de HS construído no grupo de Optometria da Universidade de Indiana (EUA). Na Figura 6.1 são destacados os principais componentes desse sistema, como por exemplo, a câmera utilizada no sistema de alinhamento do feixe *laser* e o sensor de HS no qual as imagens são formadas.

Figura 6.1 - Arquitetura do sistema de aquisição baseado na técnica de HS [Vis01].

Foram adquiridas imagens dos olhos de 100 pacientes, cujas pupilas foram dilatadas e a acomodação paralisada com a aplicação de ciclopentalato (0.5%, uma gota), sendo três imagens do olho direito (OD) e três do olho esquerdo (OS), totalizando seis imagens por paciente (600 exemplos) [Thi02b]. Cada imagem é identificada por um rótulo em que consta o número do paciente, o olho (direito ou esquerdo) e o número da aquisição, por exemplo, 24o64_OD1. Cada imagem é associada a três medidas, esférica (S), cilíndrica (C) e eixo do astigmatismo (A), que caracterizam os vícios refrativos do olho correspondente. Na Figura 6.2 é exibida a relação entre as medidas esférica (S), cilíndrica (C) e eixo do astigmatismo (A) e os vícios de refração ocular, miopia, hipermetropia e astigmatismo. Uma medida S negativa indica que o olho diagnosticado sofre de miopia (dificuldade de enxergar à distância), e quando positiva indica a presença de hipermetropia (dificuldade de enxergar de perto). As medidas C e A estão

relacionadas ao astigmatismo e ao eixo do astigmatismo, respectivamente (visão distorcida, tanto de longe quanto perto).

Figura 6.2 – Relação entre as informações que permitem a medição dos vícios refrativos.

Essas medidas são utilizadas no treinamento dos algoritmos de AM, e foram obtidas pelo algoritmo desenvolvido por pesquisadores da equipe de Optometria da Universidade de Indiana [Hon02], o qual é baseado no princípio do polinômio de Zernike [Lia94] descrito na Seção 3.4. As medidas associadas às imagens do banco têm resolução de 0.25 D para *S* e *C*, com valores para esférica (*S*) variando no intervalo entre -1.75 D e +0.25 D, e com valores para cilíndrica (*C*) variando entre 0.0 D e 1.25 D; valores do eixo do astigmatismo (*A*) variam entre 0° e 180°, com resolução de 5°. O banco é constituído de imagens em tons de cinza (*grayscale*) com resolução de 536 x 414 *pixels*.

6.2 Pré-processamento

O pré-processamento das imagens oftalmológicas engloba duas fases. A primeira consiste em centralizar o círculo de HS, que corresponde ao diâmetro da pupila e à região de interesse para a análise, em uma nova imagem de tamanho padrão. Esse

tamanho padrão foi especificado de acordo com o maior diâmetro de pupila (maior tamanho do círculo de HS) encontrado no banco de imagens. A segunda etapa consiste na eliminação de ruídos devidos às imperfeições e desalinhamento óptico no equipamento de aquisição. Na Figura 6.3 é apresentada uma visão geral do processamento efetuado por esse módulo.

Figura 6.3 – Operações efetuadas pelo módulo de pré-processamento das imagens oftalmológicas.

6.2.1 Centralização e padronização do círculo de Hartmann-Shack

A informação referente aos vícios refrativos está contida no círculo de HS da imagem. É necessário identificar onde está esse círculo, que corresponde à Região de Interesse (ROI – *Region Of Interest*) em cada imagem, e enquadrá-lo de forma que o mesmo ocupe a região central de uma nova imagem de tamanho padrão; esse processo é chamado de “*crop*”. Além de efetuar a padronização necessária para o processo de extração de características, o “*crop*” reduz as imagens para 400 x 400 *pixels*.

Para descrever o processamento efetuado é adotada a seguinte terminologia, ilustrada na Figura 6.4.

- **Pontos:** são regiões correspondentes à reflexão da frente de onda de luz no sensor de HS. Um *ponto* é uma área de 5×5 *pixels* com alta intensidade luminosa. Todos os *pontos* estão no interior do círculo de HS (Figura 6.3);
- **Fronteira:** região delimitada pela segmentação em torno do *ponto* ou dos *pontos*;
- **Pontos extremos:** os pontos mais externos, que são utilizados para delimitar a região correspondente ao círculo de HS.

Figura 6.4 – Imagem de HS apresentando os elementos da imagem referenciados na descrição dos algoritmos de pré-processamento.

Para delimitar o círculo de HS é preciso identificar os “*pontos*” mais extremos nas linhas e colunas da grade de pontos que define o círculo. Isso é feito automaticamente por meio de uma análise do histograma da imagem que permite identificar o valor do *pixel* limiar que separa os “*pontos*” do fundo preto da imagem. Contudo, para que o algoritmo de identificação do limiar tivesse um bom desempenho foi necessário classificar as imagens em “claras” e “escuras” (Figura 6.5), pois cada um dos grupos de imagens apresenta um valor de limiar diferente.

Figura 6.5 – Exemplos de imagens do tipo “clara” e “escura”.

O cálculo do valor de limiar é feito a partir do histograma, verificando qual o valor de *pixel* para o qual a quantidade acumulada de *pixels*, computada a partir dos *pixels* de intensidade zero, corresponde a 90% do total dos *pixels* da imagem (Figura 6.6). Completando os 90%, a intensidade do último *pixel* computado na somatória corresponde ao limiar buscado. Empiricamente foi adotada a intensidade 20 como o separador entre as imagens “claras” e “escuras”, isto é, se a somatória acumulada da quantidade de *pixels* em cada intensidade não atingiu 90% do total dos *pixels* antes da intensidade 20 no histograma, o algoritmo adota o valor 80% para calcular o limiar (imagem “clara”).

Com a ROI delimitada por retas horizontais e verticais traçadas com base no limiar obtido (Figura 6.7b) são calculadas as coordenadas (x , y) do *pixel* central (Figura 6.7c). A partir do *pixel* central são contados 200 *pixels* para direita e para a esquerda, e para cima e para baixo, de forma que o centro da ROI ao centro da nova imagem de 400 x 400 *pixels*. Caso a ROI ultrapasse as fronteiras da imagem original, como ilustrado na Figura 6.7e, os *pixels* são preenchidos com valor médio dos *pixels* de fundo na ROI. A Figura 6.7d apresenta a nova imagem obtida ao final do processo.

Figura 6.6 – Ilustração do cálculo do limiar.

Figura 6.7 – Sequência de passos do processo de centralização e padronização da imagem. (a) imagem inicial; (b) e (e) definição da região de interesse; (c) localização do centro da ROI; e (d) resultado do processo.

6.2.2 Eliminação de ruídos

Várias imagens apresentam ruídos e características que podem distorcer os resultados do processo de extração de características, como ilustrado na Figura 6.8. Foi implementado um processo de segmentação específico que considera as características geométricas dos “*pontos*”, com base na análise do histograma da imagem.

Figura 6.8 – Imagem com ruído oriundo do processo de aquisição.

O processo de segmentação se inicia com a delimitação das regiões próximas a cada um dos “*pontos*” da imagem, criando as chamadas “*fronteiras*” (Figura 6.4). Contudo, nesse processo, devido a problemas de aquisição múltiplos “*pontos*” podem ficar no interior de uma mesma “*fronteira*”, isto é, não são delimitados apropriadamente e não são separados dos seus vizinhos. Esse problema é ilustrado na imagem da Figura 6.9b, em que há uma região delimitada na qual os “*pontos*” não foram corretamente separados. Essa região precisa ser processada de forma a separar os “*pontos*”, o que é feito decompondo a região. A decomposição é um processo semi-automático iniciado pelo usuário, que seleciona o centro dos “*pontos*” contidos no interior da “*fronteira*”. A partir de cada centro selecionado é iniciado um processo de crescimento de região que identifica os “*pontos*” independentes. Na Figura 6.9c é exibido um exemplo de imagem com três “*pontos*” selecionados dentro de uma “*fronteira*”. Esse processo de crescimento a partir de cada centro dos “*pontos*” selecionados percorre três regiões

diferentes, determinadas por máscaras concêntricas com tamanhos pré-estabelecidos. Ao término do processo os *pixels* da região interna da antiga “*fronteira*” que não pertencem a nenhum dos novos “*pontos*” recebem valores equivalentes à média dos valores dos *pixels* da “*fronteira*” que os continha. Os valores dos *pixels* para o preenchimento das duas regiões mais externas são calculados com base na média dos *pixels* de cada “*fronteira*” da vizinhança, permitindo que as novas “*fronteiras*” mantenham características similares a da sua vizinhança. A Figura 6.9d apresenta a imagem obtida ao final do processo, com os “*pontos*” devidamente separados.

Figura 6.9 – Sequência de passos do processo de eliminação de ruído. (a) imagem inicial; (b) segmentação para localização das “*fronteiras*”; (c) selecionando o centro dos três “*pontos*” para realização da decomposição da região; e (d) resultado do processo de decomposição das regiões.

6.3 Extração de características

Como comentado anteriormente, a extração de características permite reduzir o número de atributos de entrada para o módulo de análise sem perder informações relevantes. Neste trabalho foram aplicadas duas técnicas de dois métodos diferentes para a extração de vetores de características das imagens da base. Da análise multivariada foi implementada a técnica da PCA (*Principal Component Analysis*), e do método de transformada de *wavelets* foi implementada a transformada de Gabor. A utilização de ambas as técnicas teve como objetivo verificar o impacto do tipo de extração no desempenho do módulo de análise. Com a PCA foram montados quatro vetores de características distintos, com diferentes números de atributos, e com a transformada de Gabor foi gerado um único vetor.

Foram implementados aplicativos para gerar automaticamente o vetor de características de cada uma das imagens da base original, processando toda a base sem intervenção do usuário. Esses aplicativos recebem como entrada os vetores de *pixels* que representam as imagens de HS já pré-processadas, e gera como saídas os vetores de características. O tempo de processamento para extrair as características de todas as imagens da base utilizando a PCA foi de aproximadamente três horas, uma média de 25 segundos por extração. Para a transformada de Gabor o tempo médio foi de aproximadamente 4.15 minutos por imagem, totalizando quase 27 horas de processamento. O *hardware* utilizado foi um Pentium III 500MHz com 512MB de memória RAM.

6.3.1 Vetores de características utilizando PCA

Inicialmente a PCA foi implementada com um algoritmo para a geração dos autovetores e autovalores baseado no método de Jacob [Wil92]. Contudo, o método mostrou-se muito lento, o processo consumia aproximadamente 100 segundos para uma imagem de 400 x 400 (160.000) *pixels* no *hardware* utilizado. Buscando maior eficiência computacional, foram adotados o método Householder e o algoritmo QL para calcular os autovetores e autovalores [Wil92]. O ganho foi de quatro vezes, isto é, a mesma imagem passou a ser processada em apenas 25 segundos.

A PCA foi aplicada nas colunas de cada imagem proveniente do pré-processamento, como ilustrado na Figura 6.10. Cada imagem foi reduzida para um vetor de características contendo 16 x 400 (6.400) atributos que mantém em torno de 55% das informações mais relevantes da imagem original. Depois de processar todo o banco de imagens com o algoritmo da PCA verificou-se que, para manter aproximadamente 55% do conteúdo relevante de todas as imagens, deveriam ser mantidas aproximadamente as 16 “colunas” mais significativas do resultado da PCA.

Com o intuito de analisar a influência da entrada reduzida pela PCA no desempenho do módulo de análise, foram geradas outras configurações de entrada variando a quantidade de colunas significativas mantidas na geração do vetor de características. Foram montados vetores de características mantendo 35 colunas (14.000 atributos), que representam em torno de 68% das informações mais relevantes da imagem original; e mantendo 70 colunas (28.000 atributos), que representam em torno de 81% das informações mais relevantes. Todos os vetores de características são normalizados antes de serem disponibilizados para o módulo de análise.

Figura 6.10 – Aplicação da PCA nas colunas da imagem pré-processada.

Outra variação na configuração do vetor de características consistiu em aplicar a PCA em duas etapas, primeiro nas colunas e depois nas linhas de cada imagem, como ilustrado na Figura 6.11. Ao final desse processo cada imagem é reduzida para um vetor de 140 x 10 (1.400) atributos. Na primeira etapa as 400 colunas da imagem original são reduzidas para 140 colunas, que representam em torno de 90% das informações mais relevantes da imagem original. O resultado é uma nova matriz de 140 x 400 (56.000) atributos. Na segunda etapa a PCA é aplicada às linhas dessa nova matriz, criando uma

outra matriz de 140 x 10 (1.400) atributos. Novamente, a decisão de manter as 10 ‘linhas’ mais significativas se deve a uma avaliação do banco de imagens que indicou que para manter 90% do conteúdo relevante da imagem já reduzida pela PCA na primeira etapa seria necessário manter um mínimo de 10 linhas. A diferença no número de colunas e linhas necessárias para manter 90% das informações mais relevantes dos dados entre a primeira e a segunda PCA é devida aos diferentes tamanhos de entrada apresentada à primeira (160.000) e à segunda (56.000) PCAs. Além disso, a entrada da segunda PCA já foi manipulada pela técnica.

Figura 6.11 – Aplicação da PCA às imagens em duas etapas.

6.3.2 Vetores de características utilizando Gabor

Como alternativa à PCA para redução do tamanho da entrada foi utilizada a transformada Wavelet de Gabor. Os dois parâmetros de entrada do processo de Gabor, *orientação* e *escala*, permitem determinar o número de características desejadas segundo a expressão $[orientação * escala * 2]$. Foram definidos empiricamente um valor de *orientação* igual a 10 e um valor de *escala* igual a 10, o que produz um vetor com 200 características relevantes da imagem. O algoritmo de Gabor implementado foi baseado no utilizado por Castanón [Cas03] para recuperação de imagens médicas por conteúdo, o qual por sua vez foi inspirado no algoritmo desenvolvido por Daugman [Dau02] para uma conhecida aplicação de reconhecimento de pessoas pela íris (biometria ocular). Os 200 atributos fornecidos para o módulo de análise foram previamente normalizados.

6.3.3 Comparação entre os vetores de características

Lembrando que o banco de imagens possui três imagens de cada olho, para comparar os vetores de características gerados para diferentes imagens do mesmo olho foi calculada a Distância Euclidiana entre os vetores gerados para essas imagens. A Distância Euclidiana (DE) entre dois vetores P e Q (gerados por uma mesma técnica) é dada por $\|P - Q\|$, segundo a equação (6.1). Se $P = (x_1, \dots, x_n)$ e $Q = (y_1, \dots, y_n)$, então:

$$DE(P, Q) = \|P - Q\| = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \dots + (x_n - y_n)^2} \quad (6.1)$$

A título de exemplo são apresentadas as Distâncias Euclidianas relativas aos vetores associados às três imagens do olho do paciente 24o68_OD (24o68_OD1; 24o68_OD2; 24o68_OD3), os quais foram agrupados em três pares (OD1 e OD2; OD2 e OD3; OD3 e OD1) e comparados dois a dois tanto para a PCA (Tabela 6.1), quanto para a Gabor (Tabela 6.2). Também foi calculada a diferença, em módulo, entre cada atributo dos vetores de características referentes aos pares de imagens acima, para traçar alguns gráficos do comportamento. Os gráficos referentes aos pares de imagens 1 e 2, por PCA e por Gabor, são apresentados nas Figuras 6.12 e 6.13. Os conjunto completo de gráficos referentes aos três pares de imagens é apresentado no Anexo 1.

Tabela 6.1 – Distâncias Euclidianas dos vetores de características obtidos com PCA.

Par de imagens	Distância Euclidiana
24o68_OD 1 e 2	3158.90
24o68_OD 2 e 3	3844.60
24o68_OD 3 e 1	3689.96

Tabela 6.2 – Distâncias Euclidianas dos vetores de características obtidos por Gabor.

Par de imagens	Distância Euclidiana
24o68_OD 1 e 2	3.34
24o68_OD 2 e 3	8.12
24o68_OD 3 e 1	9.65

Figura 6.12 – Diferença entre os vetores de características obtidos com PCA – imagens 1 e 2.

Figura 6.13 – Diferença entre os vetores de características obtidos por Gabor – imagens 1 e 2.

Tanto as Distâncias Euclidianas apresentadas nas tabelas como os gráficos de comportamento mostram que imagens de um mesmo olho apresentam diferenças significativas entre si, o que prejudica a formação de um padrão de reconhecimento e dificulta uma generalização pelo módulo de análise. O número de exemplos de imagens do olho de um mesmo paciente é pequeno para realizar um estudo em profundidade sobre os dados extraídos e as relações entre os mesmos. Um número maior de exemplos de cada olho (por exemplo, 30 exemplos) seria necessário para um melhor desempenho do processo de aprendizado, como discutido no Capítulo 7.

Capítulo 7

---

7. RECONHECIMENTO E ANÁLISE

Este capítulo relata a concepção e implementação do módulo de análise dos vícios refrativos. Esse módulo implementa a Etapa 3 do sistema de medição, recebendo como entrada o vetor de características de uma imagem, gerado no módulo de extração de características (Etapa 2), e produzindo como saída estimativas dos valores das três medidas (S , C e A) associadas à imagem correspondente. São apresentadas várias alternativas investigadas para a implementação do módulo de análise, buscando identificar a mais adequada para o sistema de medição, e os resultados obtidos com cada uma delas. Inúmeros experimentos e testes utilizando diferentes modelos e arquiteturas de RNAs foram realizados e proporcionaram um aprimoramento técnico gradual que permitiu chegar nas alternativas de implementação descritas neste capítulo, que são as que apresentaram melhores resultados.

Na Seção 7.1 deste capítulo são discutidas as limitações do banco de imagens e suas consequências para o desempenho do módulo de análise. Na Seção 7.2 são descritos os experimentos realizados com o algoritmo SVM utilizando como entrada os atributos extraídos com a PCA. Foram testadas várias configurações variando a quantidade de atributos do vetor de características. Na Seção 7.3 são descritos os experimentos realizados com a mesma técnica, SVM, utilizando os atributos extraídos pela transformada *wavelet* de Gabor. Os resultados e desempenho do módulo de análise com as duas técnicas são comparados. A Seção 7.4 apresenta uma abordagem alternativa que utiliza como entrada os mesmos atributos fornecidos pela transformada de Gabor, porém aplicando redes RBF para a análise. O objetivo foi avaliar o desempenho das SVMs em relação a uma outra técnica de AM.

O processo de aprendizado foi tratado inicialmente como um problema de regressão, em que é estimado um valor contínuo para as saídas S , C e A ; e depois como um problema de classificação, em que o espectro de medidas disponível para cada saída foi distribuído em um conjunto pré-definido de classes e a saída gerada pelo algoritmo de aprendizado é a classe estimada para cada medida. No aprendizado supervisionado, todo exemplo possui um atributo especial, denominado rótulo ou classe, que descreve o fenômeno de interesse, isto é, o conceito/meta que se deseja aprender para fazer previsões a respeito [Mon03].

7.1 Considerações sobre o banco de imagens

Observando-se as três medidas – esférica (S), cilíndrica (C) e eixo do astigmatismo (A) – associadas a cada imagem de HS que caracterizam os vícios refrativos verificou-se dois problemas que prejudicaram seriamente o treinamento do algoritmo de aprendizado no módulo de análise. O primeiro problema está relacionado à presença de medidas ambíguas de S , C e A associadas a múltiplas imagens do mesmo olho. Como discutido na Seção 6.3.3, em muitos casos existe uma grande diferença entre as três imagens do mesmo olho de um mesmo paciente, e cada uma delas está associada a medidas diferentes de S , C e A . Um exemplo do problema pode ser observado na Tabela 7.1. O olho direito (OD) do paciente “24o6p” possui as três medidas de S bastante diferentes. Essa diferença é salientada quando são utilizadas as medidas no padrão de “Auto Ref”, as quais são calculadas a partir das medidas originais do banco de imagens (Tabela 7.1).

O padrão “Auto Ref” utilizado para apresentar a saída adota a resolução da saída utilizada por vários auto-refratores comerciais, nos quais as medidas são apresentadas com resolução de 0.12 ou 0.25 Dioptrias para miopia, hipermetropia (esférica) e astigmatismo (cilíndrica); e de 5° para o eixo do astigmatismo. São essas as medidas utilizadas no treinamento dos algoritmos de AM. Para treinar adequadamente o módulo de análise seria importante que as três medidas de um mesmo olho fossem semelhantes, como ocorre com medidas “Auto Ref” de C do paciente “24o65” (Tabela 7.1). Uma alternativa, caso exista dificuldade de replicabilidade na metodologia de obtenção das medidas, seria aumentar o número de exemplos por olho (por exemplo, 30 imagens). A existência de poucos exemplos muitas vezes dificulta a geração de um bom modelo

estatístico baseado em padrões, o que diminui as chances de um desempenho adequado na fase de generalização por um sistema baseado em AM.

Tabela 7.1 – Exemplos da ambigüidade das medidas de S , C e A associadas à imagens distintas do mesmo olho.

Vícios	24o65_OS1	->	S :	-0.879487	C :	0.244773	A :	119.159743
AutRef	24o65_OS1	->	S :	-1.00	C :	0.25	A :	120
Vícios	24o65_OS2	->	S :	-0.851421	C :	0.205643	A :	102.287844
AutRef	24o65_OS2	->	S :	-0.75	C :	0.25	A :	100
Vícios	24o65_OS3	->	S :	-0.822712	C :	0.291625	A :	98.185900
AutRef	24o65_OS3	->	S :	-0.75	C :	0.25	A :	100
Vícios	24o6c_OD1	->	S :	-1.300193	C :	0.096985	A :	154.486751
AutRef	24o6c_OD1	->	S :	-1.25	C :	0.00	A :	155
Vícios	24o6c_OD2	->	S :	-1.639105	C :	0.360610	A :	88.633066
AutRef	24o6c_OD2	->	S :	-1.75	C :	0.25	A :	90
Vícios	24o6c_OD3	->	S :	-1.723763	C :	0.045326	A :	71.084681
AutRef	24o6c_OD3	->	S :	-1.75	C :	0.00	A :	70
Vícios	24o6p_OD1	->	S :	-0.561360	C :	0.557320	A :	99.204470
AutRef	24o6p_OD1	->	S :	-0.50	C :	0.50	A :	100
Vícios	24o6p_OD2	->	S :	-0.702275	C :	0.610150	A :	96.093289
AutRef	24o6p_OD2	->	S :	-0.75	C :	0.50	A :	95
Vícios	24o6p_OD3	->	S :	-0.346040	C :	0.348279	A :	102.392167
AutRef	24o6p_OD3	->	S :	-0.25	C :	0.25	A :	100

A segunda limitação do banco de imagens é consequência direta do primeiro problema, e é o ruído presente nos valores medidos. Observa-se nas medidas do paciente “24o65” (Tabela 7.1) que a diferença entre os valores apresentados para a saída S é 0.25; para C é 0.25; e para A é 20. Existem duas imagens com valores de medidas idênticos (24o65_OS2 e 24o65_OS3) e outra com uma grande variação em relação às outras duas (24o65_OS1). As diferenças nas medidas de S e C estão dentro da margem de erro permitida pela resolução do equipamento de aquisição das imagens, que para S e C é de ± 0.25 Dioptrias (D), e para A é de ± 5 graus ($^\circ$). Porém, o mesmo não ocorre com a medida de A . Em muitas das imagens do banco é possível verificar essa variação superior à margem de erro tolerada. Existem casos graves em que a diferença entre as medidas chega a ser de 0.75 D para S , 0.5 D para C e 70° para A .

Diante desse fato, foram retirados do banco de imagens original os exemplos referentes a um mesmo olho cujas medidas apresentavam diferença superior a 0.25 D

para S e C , e superior a 5° para A . O banco de imagens resultante ficou com 426 exemplos. Nas Tabelas 7.2, 7.3 e 7.4 são apresentadas as quantidades de exemplos de cada medida de S , C e A com sua respectiva distribuição por classes (em porcentagem) que integram o banco de imagens filtrado, considerando as medidas no padrão “Auto Ref”. Para esférica (S) as medidas variam entre -1.75 D e $+0.25$ D; para cilíndrica (C) variam entre 0.0 D e 1.25 D; e para o eixo do astigmatismo (A) variam entre 0° e 180° .

Contudo, feita a “limpeza” do banco de imagens observou-se que algumas das medidas de A não alcançaram a quantidade mínima de exemplos para realizar o treinamento, que foi estabelecida em cinco exemplos por medida (linhas sombreadas na Tabela 7.4). Conseqüentemente, os exemplos com essas medidas foram descartados no treinamento, mantendo-se apenas os que apresentavam a quantidade mínima. Descartados esses novos exemplos, o banco de imagens resultante para A passou a ter 397 exemplos (Tabela 7.5).

Tabela 7.2 – Quantidade de exemplos da medida S do novo banco de imagens.

Medida S no formato Aut Ref	Quantidade de exemplos	Distribuição de classes (%)
1.75	5	1,17%
1.50	13	3,05%
1.25	27	6,34%
1.00	42	9,86%
0.75	82	19,25%
0.50	115	27%
0.25	99	23,24%
0.00	36	8,45%
-0.25	7	1,64%

Tabela 7.3 – Quantidade de exemplos da medida C do novo banco de imagens.

Medida C no formato Aut Ref	Quantidade de exemplos	Distribuição de classes (%)
0.00	30	7,04%
0.25	229	53,76%
0.50	113	26,52%
0.75	31	7,28%
1.00	15	3,52%
1.25	8	1,88%

Tabela 7.4 – Quantidade de exemplos da medida A do novo banco de imagens.

Medida A no formato Aut Ref	Quantidade de exemplos	Distribuição de classes (%)
0	0	0%
5	5	1,17%
10	17	4,01%
15	7	1,64%
20	5	1,17%
25	3	0,71%
30	2	0,47%
35	9	2,11%
40	5	1,17%
45	4	0,94%
50	9	2,11%
55	7	1,64%
60	14	3,29%
65	18	4,22%
70	21	4,93%
75	30	7,04%
80	25	5,87%
85	46	10,8%
90	37	8,7%
95	19	4,46%
100	36	8,45%
105	21	4,93%
110	10	2,35%
115	9	2,11%
120	9	2,11%
125	12	2,82%
130	9	2,11%
135	1	0,23%
140	3	0,7%
145	2	0,47%
150	4	0,94%
155	5	1,17%
160	4	0,94%
165	1	0,23%
170	2	0,47%
175	12	2,82%
180	3	0,7%

Tabela 7.5 – Quantidade de exemplos da medida *A* que satisfazem o critério de número mínimo.

Medida <i>A</i> no formato Aut Ref	Quantidade de exemplos	Distribuição de classes (%)
5	5	1,26%
10	17	4,28%
15	7	1,75%
20	5	1,26%
35	9	2,27%
40	5	1,26%
50	9	2,27%
55	7	1,75%
60	14	3,53%
65	18	4,53%
70	21	5,29%
75	30	7,56%
80	25	6,3%
85	46	11,59%
90	37	9,32%
95	19	4,79%
100	36	9,07%
105	21	5,29%
110	10	2,52%
115	9	2,27%
120	9	2,27%
125	12	3,02%
130	9	2,27%
155	5	1,26%
175	12	3,02%

7.2 Experimentos utilizando PCA e SVM

Em vista das limitações do banco de imagens optou-se por investigar o algoritmo SVM para a implementação do módulo de análise. Como discutido na Seção 5.4, as SVMs possuem uma boa capacidade de generalização por seguir o princípio de minimização do risco estrutural, ao invés do princípio de minimização do risco empírico empregado por RNAs convencionais. A minimização do risco estrutural reduz o erro de generalização, enquanto o risco empírico minimiza o erro de treinamento [Vap95].

O módulo de análise utilizando SVM foi implementado baseando-se na ferramenta *Torch*, (linguagem C) para a plataforma LINUX [Tor02]. Nos experimentos foram manipulados os seguintes parâmetros do modelo SVM: o tipo de *kernel* (linear, polinomial, Gaussiano e sigmoidal), o critério de parada e os parâmetros dos *kernels*. Foi desenvolvido um algoritmo para gerar automaticamente diferentes combinações de

valores possíveis para o conjunto de parâmetros, inclusive ajustando os parâmetros específicos dos *kernels* (Tabela 7.6) e, posteriormente, treinar o modelo resultante com o sub-conjunto de treinamento e validação. Outro programa obtém a medida de desempenho (estimativa de erro) para o modelo definido pelo conjunto de parâmetros sendo testado, utilizando o sub-conjunto de teste. O conjunto de parâmetros do modelo que apresentou a melhor medida de desempenho foi armazenado para posterior utilização no módulo de análise. Devido ao grande número de combinações possíveis para os parâmetros, os treinamentos e a verificação do melhor modelo demandaram um longo tempo e esforço.

Tabela 7.6 – Tipos de kernel para modelos de SVM [Gun98].

Tipo de <i>kernel</i>	Função
linear	$K(x, y) = (x \cdot y)$
polinomial	$K(x, y) = ((x \cdot y) + 1)^d$
gaussiano	$K(x, y) = \exp\left(-\frac{(x - y)^2}{2\sigma^2}\right)$
sigmoidal	$K(x, y) = \exp\left(-\frac{ x - y }{2\sigma^2}\right)$

Para estimar o erro foi adotado o método de reamostragem aleatória (*random resampling*) [Car99], isto é, o conjunto de dados original com 426 exemplos foi particionado 10 vezes em dois subconjuntos, um sub-conjunto de treinamento e validação com 80% dos exemplos, e um sub-conjunto de teste com 20% dos exemplos. Essas 10 partições foram montadas aleatoriamente e têm como objetivo aumentar a significância estatística dos resultados. Na montagem dos subconjuntos procurou-se manter uma distribuição uniforme dos exemplos de cada medida de *S*, *C* e *A*.

Quando o aprendizado é tratado como um problema de classificação, uma medida de desempenho comumente usada é a taxa de erro do classificador. Usualmente, a taxa de erro é obtida utilizando a equação (7.1), a qual compara a classe verdadeira de cada exemplo com a classe atribuída pelo classificador *h*. O operador $\|E\|$ retorna 1 se a expressão *E* for verdadeira, e zero caso contrário; *n* é o número de exemplos. Para problemas de regressão, o erro pode ser estimado calculando-se a distância entre o valor real e o valor atribuído pelo classificador. Usualmente, a medida adotada é o erro quadrático médio (*mean squared error* – MSE) [Mon03].

$$err(h) = \frac{1}{n} \sum_{i=1}^n \| y_i - h(x_i) \| \quad (7.1)$$

Dado um algoritmo A e um conjunto de exemplos T , assume-se que T seja dividido em r partições. Para cada partição i , é induzida a hipótese h_i e calculado o erro denotado por $err(h_i)$, $i = 1, 2, \dots, r$. Em seguida são calculados a média, a variância e o desvio padrão para todas as partições utilizando as equações (7.2), (7.3) e (7.4), respectivamente [Mon03].

$$mean(A) = \frac{1}{r} \sum_{i=1}^r err(h_i) \quad (7.2)$$

$$var(A) = \frac{1}{r} \left[\frac{1}{r-1} \sum_{i=1}^r (err(h_i) - mean(A))^2 \right] \quad (7.3)$$

$$sd(A) = \sqrt{var(A)} \quad (7.4)$$

A arquitetura adotada no experimento é apresentada na Figura 7.1. Para cada uma das medidas S , C e A foi treinado um modelo de SVM. É possível observar na Figura que, inicialmente, o processo de aprendizado foi tratado como um problema de regressão, isto é, foram utilizadas SVMs para regressão e as saídas S , C e A consistem, cada uma, em um valor contínuo estimado para cada uma das medidas. Por exemplo, a SVM para estimar S possui como entrada um vetor de características de 6400 atributos extraídos utilizando PCA e gera como saída um valor real. Posteriormente, este valor é convertido no padrão de medida “Auto Ref”, que apresenta o valor segundo a resolução de um auto-refrator comercial, de 0,25 D.

Figura 7.1 – Módulo de análise proposto inicialmente com medidas regressivas.

Os erros médios obtidos com os conjuntos de teste definidos pelas 10 partições pelas melhores configurações de SVM, para cada uma das medidas S , C e A são mostrados na Tabela 7.7. A segunda coluna da tabela apresenta a quantidade de exemplos de cada uma das medidas presentes em cada conjunto de teste. Na coluna com o título “Medida exata” são apresentadas as taxas médias (em porcentagem) e o desvio padrão (em porcentagem) do erro das SVMs para as medidas de S , C e A . A coluna com o título “Propagação do erro” também apresenta o erro médio e desvio padrão, porém considerando uma tolerância do erro das medidas de S e C de ± 0.25 D, e uma tolerância do erro de A de $\pm 5^\circ$. Isso significa que, se uma medida S ou C possui valor igual a 0.5 D, são computados como acertos do classificador os valores 0.5 D, 0.25 D ou 0.75 D. Para uma medida A cujo valor exato é 30° são computados como acertos os valores 30° , 25° ou 35° .

Tabela 7.7 – Desempenho das melhores configurações de SVM para a arquitetura regressiva.

	Total de exemplos	Medida exata		Propagação do erro	
		Erro Médio	Desvio Padrão	Erro Médio	Desvio Padrão
S	82	78.05%	$\pm 1.71\%$	39.03%	$\pm 1.92\%$
C	83	73.49%	$\pm 0.82\%$	18.07%	$\pm 0.93\%$
A	70	94.29%	$\pm 2.51\%$	81.43%	$\pm 2.33\%$

Alternativamente a uma arquitetura para regressão, decidiu-se investigar se uma arquitetura de SVM para classificação (Figura 7.2) apresentaria melhor desempenho, e o módulo de análise foi alterado para tratar o problema de apreendizado como um problema de classificação. Com isso, ao invés de aproximar um valor de saída as SVMs devem associar uma imagem a uma das classes de saída. O conjunto de medidas de *S* foi separado em 9 classes, correspondentes a cada grupo de medidas entre -1.75 D e +0.25 D, com intervalos de 0.25 D, indicado na Tabela 7.2; o conjunto de medidas *C* foi separado em 6 classes, correspondentes a cada grupo de medidas entre 0.0 D e 1.25 D, com intervalos de 0.25 D, indicado na Tabela 7.3; e o conjunto de medidas de *A* foi separado em 25 classes, segundo a distribuição das medidas de *A* apresentada na Tabela 7.5.

As saídas de *S*, *C* e *A* são representadas por vetores de 9, 6 ou 25 elementos, respectivamente, em que cada elemento do vetor corresponde a uma classe na posição correspondente. Por exemplo, a posição 1 do vetor está associada à primeira classe das medidas, a posição 2 à segunda classe das medidas, e assim por diante. Assim, uma SVM para medir *S* possui como entrada um vetor de 6400 atributos e como saída um vetor de 9 elementos: a posição do maior dos 9 elementos do vetor de saída indica a correspondente classe da saída. Os resultados obtidos com as melhores configurações da arquitetura de classificação, para cada uma das medidas, são mostrados na Tabela 7.8.

Tabela 7.8 – Desempenho das melhores configurações de SVM para a arquitetura de classificação.

	Total de exemplos	Medida exata		Propagação do erro	
		Erro Médio	Desvio Padrão	Erro Médio	Desvio Padrão
<i>S</i>	82	75.61%	± 1.01%	30.49%	± 1.06%
<i>C</i>	83	45.78%	± 0.81%	12.05%	± 0.83%
<i>A</i>	70	87.14%	± 1.62%	74.29%	± 1.63%

Figura 7.2 – Módulo de análise para classificação.

7.2.1 Variando a quantidade de atributos de entrada

Os resultados obtidos com a transformação do problema de regressão em um problema de classificação apresentaram uma melhora, porém insuficiente para garantir um bom desempenho do sistema de medição. Optou-se, então, por aumentar a quantidade de atributos do vetor de características fornecido como entrada. Foram testadas três variações, com 1400 (resultados de erro médio e desvio padrão na Tabela 7.9), 14000 (resultados na Tabela 7.10) e 28000 atributos (resultados na Tabela 7.11). Como pode ser observado, os resultados apresentados foram inferiores ou semelhantes aos obtidos utilizando-se 6400 atributos. Por isso, foi mantida a configuração inicial (6400 atributos).

Tabela 7.9 – Resultados da arquitetura de classificação com 1400 atributos.

	Total de exemplos	Medida exata		Propagação do erro	
		Erro Médio	Desvio Padrão	Erro Médio	Desvio Padrão
<i>S</i>	82	78.05%	± 1.51%	40.24%	± 1.05%
<i>C</i>	83	45.78%	± 0.81%	13.25%	± 0.82%
<i>A</i>	70	88.57%	± 1.61%	75.71%	± 1.62%

Tabela 7.10 – Resultados da arquitetura de classificação com 14000 atributos.

	Total de exemplos	Medida exata		Propagação do erro	
		Erro Médio	Desvio Padrão	Erro Médio	Desvio Padrão
<i>S</i>	82	77.39%	± 1.60%	30.49%	± 1.57%
<i>C</i>	83	46.99%	± 0.90%	13.25%	± 0.88%
<i>A</i>	70	88.57%	± 1.72%	80.00%	± 1.70%

Tabela 7.11 – Resultados da arquitetura de classificação com 28000 atributos.

	Total de exemplos	Medida exata		Propagação do erro	
		Erro Médio	Desvio Padrão	Erro Médio	Desvio Padrão
<i>S</i>	82	78.05%	± 1.00%	32.93%	± 1.07%
<i>C</i>	83	46.99%	± 1.10%	14.46%	± 1.20%
<i>A</i>	70	92.86%	± 1.41%	84.29%	± 1.44%

7.2.2 Balanceamento artificial do banco de imagens

Com o intuito de verificar quanto os resultados obtidos são influenciados pelo desbalanceamento das classes, i.e., pelas diferentes quantidades de exemplos por classe, buscou-se balancear a quantidade de exemplos de cada classe de medidas antes de apresentá-los ao módulo de análise. Foram montadas 10 novas partições dos exemplos de entrada, contendo dois conjuntos cada: um conjunto de treinamento e validação, com 80% dos exemplos, e um conjunto de teste, com 20% dos exemplos. Para as medidas de *S* e *C* os conjuntos foram constituídos por classes com mais de 30 exemplos, e para as medidas de *A* os conjuntos foram constituídos por classes com mais de 25 exemplos. Foram escolhidas cinco classes para *S* (Tabela 7.12), quatro classes para *C* (Tabela 7.13) e quatro classes para *A* (Tabela 7.14). Os resultados obtidos são apresentados na Tabela 7.15.

Os resultados obtidos indicam a importância de balancear o banco de imagens para facilitar o treinamento e, conseqüentemente, obter um melhor desempenho na fase de teste, principalmente se existir um grande número de classes (observar a diferença entre os resultados da medida *A*). Este fator deve ser levado em consideração no momento da construção de um banco de imagens para teste. Contudo, observou-se que os resultados da medida *C* foram inferiores aos resultados apresentados na Tabela 7.8. Isso pode estar ocorrendo devido a duas das classes originais (medidas iguais a 0.25 e 0.50, na Tabela 7.8) possuírem uma quantidade de exemplos muito elevada em comparação às demais

classes. O alto índice de acerto é decorrente da maior quantidade de exemplos dessas classes nos conjuntos de teste.

Tabela 7.12 – Balanceamento artificial da medida *S* – 5 classes (150 exemplos).

Medida <i>S</i> no formato Aut Ref	Quantidade de exemplos	Exemplos balanceados
1.00	42	30
0.75	82	30
0.50	115	30
0.25	99	30
0.00	36	30

Tabela 7.13 – Balanceamento artificial da medida *C* – 4 classes (120 exemplos).

Medida <i>C</i> no formato Aut Ref	Quantidade de exemplos	Exemplos balanceados
0.00	30	30
0.25	229	30
0.50	113	30
0.75	31	30

Tabela 7.14 – Balanceamento artificial da medida *A* – 4 classes (100 exemplos).

Medida <i>A</i> no formato Aut Ref	Quantidade de exemplos	Exemplos balanceados
75	30	25
80	25	25
85	46	25
90	37	25

Tabela 7.15 – Resultados da arquitetura de classificação com balanceamento artificial (6400 atributos).

	Total de exemplos	Medida exata		Propagação do erro	
		Erro Médio	Desvio Padrão	Erro Médio	Desvio Padrão
<i>S</i>	82	63.33%	± 1.70%	26.67%	± 1.77%
<i>C</i>	83	50.00%	± 2.00%	25.00%	± 2.10%
<i>A</i>	70	65.00%	± 2.30%	10.00%	± 2.34%

7.3 Experimentos utilizando Gabor e SVM

Nessa seção são apresentados os resultados dos experimentos realizados com os vetores de atributos das imagens extraídos pela técnica de transformada *wavelet* de Gabor, ao invés da técnica PCA. O intuito foi verificar qual das duas técnicas de

extração possibilitaria um melhor desempenho do módulo de análise. A Tabela 7.16 apresenta os resultados obtidos. A comparação desses resultados com os apresentados na Tabela 7.8, referentes à técnica PCA, sugere que a transformada de Gabor pode ser a melhor opção para o módulo de análise. A configuração do modelo SVM que apresentou melhor desempenho para as medidas S e A utilizou o *kernel* polinomial (parâmetros para S , x igual a 120, y igual a 38 e d igual a 2 e para A , x igual a 53, y igual a 9 e d igual a 3) e para a medida C foi a que utilizou o *kernel* Gaussiano com os parâmetros x igual a 210, y igual a 9 e σ igual a 48. Na Tabela 7.17 são apresentados os resultados dos experimentos realizados com o banco de imagens balanceado, seguindo os mesmos critérios adotados nos experimentos descritos na Subseção 7.2.2. Esses resultados podem ser comparados aos da Tabela 7.15 e, novamente, sugerem a transformada de Gabor como a melhor opção para o sistema. Nesse experimento, a configuração do modelo de SVM que obteve o melhor resultado para a medida S utilizou o *kernel* linear com os parâmetros, x igual a 25 e y igual a 130 e para as medidas C e A , o *kernel* polinomial (parâmetros para C , x igual a 105, y igual a 80 e d igual a 2 e para A , x igual a 65, y igual a 12 e d igual a 2).

Tabela 7.16 – Resultados da arquitetura de classificação utilizando Gabor (200 atributos).

	Total de exemplos	Medida exata		Propagação do erro	
		Erro Médio	Desvio Padrão	Erro Médio	Desvio Padrão
S	82	62.20%	$\pm 1.10\%$	18.29%	$\pm 1.12\%$
C	83	42.17%	$\pm 1.02\%$	9.64%	$\pm 1.00\%$
A	70	81.43%	$\pm 1.61\%$	70.00%	$\pm 1.64\%$

Tabela 7.17 – Resultados da arquitetura de classificação com balanceamento artificial utilizando Gabor.

	Total de exemplos	Medida exata		Propagação do erro	
		Erro Médio	Desvio Padrão	Erro Médio	Desvio Padrão
S	82	56.67%	$\pm 1.70\%$	13.33%	$\pm 1.64\%$
C	83	41.67%	$\pm 2.12\%$	20.83%	$\pm 2.09\%$
A	70	30.00%	$\pm 2.40\%$	0.00%	$\pm 2.12\%$

Contudo, a comparação dos dois algoritmos considerando apenas a taxa de erro na classificação não permite concluir se um algoritmo é de fato melhor do que o outro. Para estabelecer tal comparação foi aplicado um teste estatístico [Mas89] que permite comprovar, com um elevado grau de confiança, se o desempenho de um algoritmo é realmente superior ao do outro. Para decidir qual dos algoritmos apresentou melhor desempenho, com grau de confiança, por exemplo, de 95%, basta assumir um caso geral

para determinar se a diferença entre os algoritmos As e Ap é significativa ou não, assumindo uma distribuição normal do erro [Wei98]. Para isso, a média e o desvio padrão combinados do erro são calculados de acordo com as equações (7.5) e (7.6), respectivamente. A diferença absoluta em desvios padrões é calculada pela equação (7.7) [Mon03].

$$mean(As - Ap) = mean(As) - mean(Ap) \quad (7.5)$$

$$sd(As - Ap) = \sqrt{\frac{sd(As)^2 + sd(Ap)^2}{2}} \quad (7.6)$$

$$ad(As - Ap) = \frac{mean(As - Ap)}{sd(As - Ap)} \quad (7.7)$$

Estabelecendo a hipótese inicial nula $H_0: As = Ap$ e a hipótese alternativa $H_1: As \neq Ap$. Se $ad(As - Ap) > 0$ então Ap supera As ; porém, se $ad(As - Ap) \geq 2.00$ (limite da região de aceitação) então Ap supera As com grau de confiança de 95%. Por outro lado, se $ad(As - Ap) \leq 0$ então As supera Ap e se $ad(As - Ap) \leq -2.00$ então As supera Ap com grau de confiança de 95%. Os limites da região de aceitação RA: (-2.00, 2.00) para esses experimentos foram baseados na tabela de distribuição de *Student t* [Mas89].

A Tabela 7.18 apresenta os valores do erro médio calculados para os resultados dos experimentos utilizando PCA e transformada de Gabor apresentados nas Tabelas 7.8 e 7.16, respectivamente. Na Tabela 7.19 são apresentados os resultados utilizando PCA e Gabor após o balanceamento artificial dos exemplos de entrada (dados apresentados nas Tabelas 7.15 e 7.17).

Tabela 7.18 – Comparação estatística entre os algoritmos PCA e Gabor.

PCA (As) – Valores do erro médio			Gabor (Ap) – Valores do erro médio		
S	C	A	S	C	A
0.756 ± 0.010	0.458 ± 0.008	0.871 ± 0.016	0.622 ± 0.011	0.421 ± 0.010	0.814 ± 0.016

PCA e Gabor	$Ad(As - Ap)$	Certeza	Região de aceitação	Hipótese H_1
S	12.75	95%	(-2.00, 2.00)	Aceita
C	4.08	95%	(-2.00, 2.00)	Aceita
A	3.56	95%	(-2.00, 2.00)	Aceita

Tabela 7.19 – Comparação estatística entre os algoritmos PCA e Gabor com balanceamento artificial.

PCA (A_s) – Valores do erro médio			Gabor (A_p) – Valores do erro médio		
S	C	A	S	C	A
0.633 ± 0.017	0.500 ± 0.020	0.650 ± 0.023	0.566 ± 0.017	0.416 ± 0.021	0.300 ± 0.024

PCA e Gabor	$ad(A_s - A_p)$	Certeza	Região de aceitação	Hipótese H_1
S	3.94	95%	(-2.00, 2.00)	Aceita
C	4.09	95%	(-2.00, 2.00)	Aceita
A	14.89	95%	(-2.00, 2.00)	Aceita

Observando a Tabela 7.18 é possível verificar que a transformada de Gabor é superior a PCA com grau de confiança maior que 95% para as três medidas (S , C e A). O mesmo acontece para os resultados apresentados na Tabela 7.19. Dessa forma, foi adotada a solução usando transformada de Gabor para o módulo de extração de características do sistema proposto.

No Anexo 2 são apresentados os gráficos que exibem as porcentagens do número de exemplos (padrões) e do erro médio total dos conjuntos de teste (10 partições) divididos por classe, para cada uma das medidas S , C e A , sem e com a propagação do erro. São apresentados, também, os gráficos correspondentes relativos aos experimentos realizados com o banco de imagens balanceado.

7.4 Experimentos utilizando Gabor e redes RBF

Como alternativa às SVMs no módulo de análise foi proposta a utilização de RNAs, mais especificamente redes RBF para implementar o processo de aprendizado. Os resultados de saída desse módulo alternativo dão parâmetros para comparar se SVM é mais adequada para esse sistema de medição em relação a RNAs, RBF em particular. Para verificar a arquitetura da rede RBF adequada ao módulo de análise foram implementados programas baseados na linguagem *Batchman* do simulador SNNS (*Stuttgart Neural Network Simulator*) para plataforma LINUX, desenvolvido pela Universidade de Stuttgart na Alemanha e disponibilizado gratuitamente na Internet [Zel95]. Foi implementada uma rotina que utiliza a técnica de Algoritmos Genéticos (AGs) para gerar combinações otimizadas dos diversos parâmetros que definem o modelo da rede (quantidade de neurônios, determinação dos centros, função de ativação: *Gaussian*, *Inverse Multiquadratic*, *Multiquadratic*, *Thin Plate Spline*, *Cubic* e

Linear e função de inicialização) de maneira a identificar a melhor arquitetura para a rede. A rotina implementada utiliza o algoritmo proposto por Lacerda para gerar arquiteturas ótimas para redes RBF [Lac98].

Nos experimentos foi utilizado o método *cross validation* [Car99], recomendado para o teste de parada do treinamento da rede. Esse método monitora o erro em um conjunto de dados independente e encerra o treinamento quando o erro de validação começa a aumentar. Este fato é um indicador de que a rede começa a memorizar o conjunto de treinamento, tornando-se incapaz de generalizar o problema (*overfitting*) [Bis95].

Da mesma forma que os experimentos anteriores envolvendo SVM, os dados de entrada originais (426 exemplos) foram divididos em dois subconjuntos, o de treinamento e validação com 80% dos exemplos, e o de teste com 20% dos exemplos. Foram criados 10 partições, cada uma com um subconjunto de treinamento e outro de teste.

Na Tabela 7.20 são apresentados os resultados obtidos. O modelo de arquitetura que obteve o melhor desempenho de classificação para a medida *S* foi constituído de 60 neurônios na camada intermediária, para função de ativação foi utilizada *Thin Plate Spline* e definida a distância entre os centros das elipses em 1.5 unidades. Para a medida *C*, o modelo de rede foi montado com 105 neurônios na camada intermediária, foi utilizada a função Gaussiana como ativação e aplicada a distância de 1 unidade entre os centros. Por fim, o modelo de rede para a medida *A*, foi formado por 185 neurônios na camada intermediária, utilizada a função *Thin Plate Spline* para ativação e 0.5 unidade de distância entre os centros.

A comparação desses resultados com os apresentados na Tabela 7.16 sugere que a opção do uso de SVM ainda é a mais eficiente. Também foram realizados experimentos com o banco de imagens balanceado, cujos resultados podem ser vistos na Tabela 7.21 e comparados aos apresentados na Tabela 7.17. Os modelos de redes RBF que apresentaram melhor desempenho foram: para a medida *S*, 40 neurônios na camada intermediária, função de ativação *Thin Plate Spline* e 1.5 unidades para distância entre os centros. Para a medida *C* o melhor modelo utilizou 80 neurônios na camada intermediária, função de ativação Gaussiana e distância de 1.5 unidade entre os centros e, finalmente, para a medida *A* a melhor configuração utilizou 90 neurônios na camada

intermediária, função de ativação Gaussiana e unidade de distância entre os centros igual a 1.

Na Tabela 7.22 é apresentada a comparação estatística dos resultados dos algoritmos SVM e RBF, levando em consideração os resultados dos experimentos apresentados nas Tabelas 7.16 e 7.20, respectivamente. Na Tabela 7.23 são apresentados os resultados obtidos com SVM e RBF com o balanceamento artificial realizado (dados apresentados nas Tabelas 7.17 e 7.21).

Tabela 7.20 – Resultados da arquitetura de classificação utilizando Gabor e redes RBF.

	Total de exemplos	Medida exata		Propagação do erro	
		Erro Médio	Desvio Padrão	Erro Médio	Desvio Padrão
<i>S</i>	82	68.29%	± 2.10%	27.06%	± 2.08%
<i>C</i>	83	50.60%	± 1.62%	21.18%	± 1.53%
<i>A</i>	70	85.71%	± 2.20%	81.18%	± 2.10%

Tabela 7.21 – Resultados com balanceamento artificial utilizando Gabor e redes RBF.

	Total de exemplos	Medida exata		Propagação do erro	
		Erro Médio	Desvio Padrão	Erro Médio	Desvio Padrão
<i>S</i>	82	63.33%	± 2.70%	26.67%	± 2.90%
<i>C</i>	83	54.17%	± 3.11%	29.17%	± 3.53%
<i>A</i>	70	40.00%	± 2.50%	25.00%	± 2.40%

Tabela 7.22 – Resultados da comparação estatística dos algoritmos SVM e RBF.

SVM (<i>As</i>) – Valores do erro médio			RBF (<i>Ap</i>) – Valores do erro médio		
<i>S</i>	<i>C</i>	<i>A</i>	<i>S</i>	<i>C</i>	<i>A</i>
0.622 ± 0.011	0.421 ± 0.010	0.814 ± 0.016	0.682 ± 0.021	0.506 ± 0.016	0.857 ± 0.022

SVM e RBF	$ad(As - Ap)$	Certeza	Região de aceitação	Hipótese H_I
<i>S</i>	-3.66	95%	(-2.00, 2.00)	Aceita
<i>C</i>	-6.20	95%	(-2.00, 2.00)	Aceita
<i>A</i>	-2.48	95%	(-2.00, 2.00)	Aceita

Tabela 7.23 – Resultados da comparação estatística dos algoritmos SVM e RBF com balanceamento artificial.

SVM (As) – Valores do erro médio			RBF (Ap) – Valores do erro médio		
S	C	A	S	C	A
0.566 ± 0.017	0.416 ± 0.021	0.300 ± 0.024	0.633 ± 0.027	0.541 ± 0.031	0.400 ± 0.025

SVM e RBF	$ad(As - Ap)$	Certeza	Região de aceitação	Hipótese H_I
S	-3.04	95%	(-2.00, 2.00)	Aceita
C	-4.71	95%	(-2.00, 2.00)	Aceita
A	-4.04	95%	(-2.00, 2.00)	Aceita

Observando a Tabela 7.22 é possível verificar que SVM é superior à RBF com grau de confiança maior que 95% para as três medidas (S , C e A). O mesmo acontece para os resultados apresentados na Tabela 7.23. Dessa forma, foi adotada a solução usando SVM para o módulo de análise do sistema proposto.

8. CONCLUSÕES

Este capítulo está dividido em duas seções. Na Seção 8.1 são discutidas as principais contribuições e apresentadas algumas considerações sobre o trabalho desenvolvido e suas limitações. Na Seção 8.2 são apresentadas sugestões para trabalhos futuros.

8.1 Contribuições e Discussão

Este trabalho apresenta uma nova abordagem para a medição dos vícios de refração ocular que utiliza técnicas de AM para a análise de imagens oftalmológicas. É proposto um sistema de análise que explora técnicas e conceitos de Entendimento/Interpretação de Imagens (*Image Understanding*), área na qual o desenvolvimento de projetos práticos que utilizam SVMs [Byu02] e RNAs [Egm02] é recente. O processo de desenvolvimento e os resultados obtidos podem servir como um estudo de caso para os pesquisadores interessados em verificar a adequação dessas técnicas de AM para a execução de tarefas no campo do Entendimento/Interpretação da Imagem. Além disso, este trabalho estabelece as bases para a geração de conhecimento em uma área multidisciplinar (computação, física, optometria e engenharia biomédica).

Os resultados apresentados permitem afirmar que os objetivos inicialmente propostos foram atingidos com êxito. Foi efetuada a concepção técnica de um sistema computacional para a medição dos vícios refrativos a partir de imagens funcionais do olho, e seus módulos estão descritos nos Capítulos 6 e 7. O sistema de medidas apresentou o melhor desempenho quando utilizada a transformada de Gabor para a extração de vetores de características das imagens e SVM para a análise desses vetores. Entretanto, todo o desenvolvimento teve como base o banco de imagens de Hartmann-Shack obtido junto ao grupo de Optometria da Universidade de Indiana (EUA), o qual,

apesar de ter se mostrado suficiente para validar a metodologia, apresenta diversas limitações.

As amostras disponíveis são em pouca quantidade e apresentam um espectro de medidas bastante reduzido: para esférica (S) o espectro varia entre -1.75 e $+0.25$; e para cilíndrica (C) varia entre 0.0 e 1.25 (resolução 0.25 D), com o eixo (A) variando entre 5° e 180° (resolução de 5°). Dentro desse espectro existem poucos exemplos de cada classe para o treinamento. Esse foi um dos fatores responsáveis pela dificuldade apresentada pelo módulo de análise em abstrair um modelo adequado para a interpretação dos atributos. Acredita-se que um maior número de exemplos representativos distribuídos uniformemente pelo espectro de medidas possível (por exemplo, S variando entre -17.00 e $+17.00$ e C variando entre 0.0 e 17.00) permitiria melhorar o desempenho do módulo de análise. Além disso, o banco de imagens deveria ser balanceado em relação às classes de medidas, i.e., possuir um número semelhante de exemplos em cada classe.

Outro fator limitante da base que prejudicou o desempenho do módulo de análise foi a diferença constatada entre as medidas fornecidas de S , C e A relativas às três imagens do olho de um mesmo paciente. Para dar continuidade à pesquisa seria necessário adquirir um novo banco de imagens com maior número de exemplos de cada olho e com medidas que apresentem menor variação. Uma sugestão seria adquirir um banco de imagens com, pelo menos, 30 imagens HS de cada olho, e o mesmo olho ser medido por um optômetro mecânico. Dessa forma, o banco teria 30 imagens de cada olho e uma única medida de S , C e A , facilitando a fase de treinamento. Pela experiência adquirida no projeto, se este procedimento for executado para um espectro de medição mais amplo os resultados do módulo de análise provavelmente serão superiores aos obtidos.

Com relação aos vetores de características fornecidos como entrada para o módulo de análise, observou-se que uma mudança na técnica adotada no módulo de extração de características (como ocorreu da PCA para Gabor) requer que seja refeito todo o processo de busca por uma configuração adequada. Seja para SVM ou RNA os resultados obtidos com a técnica anterior não servem de referência para a nova situação, o que dificulta uma decisão de mudança da técnica de extração de características. Outro fator crítico foi o tempo da fase de treinamento utilizando AGs para manipular os parâmetros das redes RBF. Por executar uma combinação dos diversos parâmetros sendo investigados, o tempo de processamento triplicou em relação ao treinamento sem

AGs. Os treinamentos para as redes RBF otimizados com AG duraram em média de 2 a 5 dias (48 horas a 120 horas) na plataforma LINUX (o *hardware* utilizado foi um Pentium III 500MHz com 512MB de memória RAM).

O fato da arquitetura do módulo de análise usando regressão não ter apresentado um desempenho adequado para o conjunto de imagens testadas não invalida a sua utilização para um sistema de medição baseado em imagens de HS. Provavelmente, o número insuficiente de exemplos representativos no espectro, para gerar um modelo correto para generalização foi decisivo para o baixo desempenho. Também foi decisiva, como salientado anteriormente, a diferença entre as medidas de S , C e A relativas às três imagens de cada olho. Apesar de todas as limitações do banco de imagens para utilização por técnicas de AM, já que ele não foi criado levando em consideração pontos importantes como balanceamento, quantidade de exemplos, e ausência de ambigüidade, o comportamento tanto das SVMs quanto das RNAs demonstrou que as mesmas conseguem responder bem às adversidades de um banco de imagens obtido por um sistema real. Na área médica é muito complicado obter um espectro adequado no qual todas as classes estão representadas em quantidade suficiente para permitir um desempenho satisfatório no módulo de análise.

Observando a facilidade das SVMs e RNAs em relacionar entradas com saídas, é possível vislumbrar uma oportunidade promissora no momento em que se deseja reproduzir, em um sistema computacional de baixo custo, resultados obtidos com um sistema complexo e custoso. Seria possível, por exemplo, detectar e diagnosticar uma doença ocular utilizando uma metodologia ou sistema complexo e depois utilizar algoritmos de AM para relacionar imagens ou vetores de informações vindos de sensores específicos a esta doença. O mesmo arcabouço básico desenvolvido para a medição de vícios de refração ocular seria aplicável à detecção de outras doenças do olho presentes em imagens HS.

8.2 Sugestões para trabalhos futuros

Por se tratar de um trabalho pioneiro, existem muitas oportunidades de continuidade na área de atuação desta tese. Um dos primeiros trabalhos sugeridos seria a investigação de outras técnicas de extração de características. Como o objetivo da extração de características é diminuir os atributos de entrada para o módulo de análise,

poderia ser testada uma compressão de imagens baseada em transformadas de *wavelet* (*Wavelet compression*), como Daubechies [Dau90] [Dau92], que é um dos métodos mais eficientes de compressão de imagens [Wav02a] [Wav02b] [Wav02c]. Seria interessante verificar o comportamento do módulo de análise com os atributos fornecidos por essa técnica e comparar os resultados aos obtidos com a PCA e com a transformada de Gabor. Um outro trabalho interessante está vinculado ao estudo das relações entre os vetores de características gerados pelas técnicas aplicadas e as medidas (no caso do projeto, os valores de S , C e A) utilizadas para supervisionar o treinamento dos algoritmos de AM. Entender esta relação pode ser fundamental no momento em que se estiver buscando extrair outros tipos de informações das imagens de HS.

Outro ponto a ser pesquisado é a influência do diâmetro do olho na imagem de HS analisada. Ele poderia ser utilizado como uma medida de supervisão adicional do módulo de análise, juntamente com S , C e A . É interessante saber se essa informação é significativa no processo de reconhecimento dos vícios refrativos, ou mesmo na análise de outros problemas oculares.

Para evitar a dependência, no momento da análise das imagens, com relação a alguma característica imposta pelo treinamento supervisionado, como é o caso da resolução das medidas de miopia, hipermetropia e astigmatismo, sugere-se utilizar um módulo de análise baseado em RNAs não supervisionadas (redes *Self-Organising*). Dessa forma, a própria rede trataria de identificar as classes utilizando características encontradas na própria imagem.

A utilização dessa metodologia de medição utilizando SVMs ou RNAs permite vislumbrar a possibilidade de análise de imagens oftalmológicas provenientes de outras técnicas de imageamento, como *Fluorescein*, *Retro-illumination* e *Optical scatter map* [Thi02a], para identificar outros problemas oculares. No caso de imagens adquiridas pela técnica de HS, acredita-se que a abordagem adotada é aplicável para medir e avaliar outras enfermidades do olho humano, como é o caso do “*Tears Film Breakup*” [Him01], o que é inviável utilizando a técnica convencional para avaliação das imagens de HS baseada em polinômios de Zernike, discutida na Seção 3.4 [Lia94].

Tendo um banco de imagens adequado, outros modelos de RNAs supervisionadas poderiam ser considerados para o sistema de medição, como GFF (*Generalized feedforward networks*), *Jordan e Elman networks* e *Time Lagged Recurrent Networks* (TLRN) e, em especial, as redes MNNs (*Modular Neural Network*) [Ned02]. As MNNs

são uma classe especial das redes MLPs que processam seus dados de entrada usando várias MLPs em paralelo e combinando os resultados de saída, gerando uma função especializada em cada sub-módulo da topologia. Diferente das MLPs, essas redes não possuem todos os neurônios internos conectados (*full interconnection*) entre as camadas. Portanto, um número menor de pesos é requisitado para o mesmo tamanho de rede, o que tende a acelerar o treinamento [Ned02].

Outra sugestão seria aplicar uma combinação de classificadores baseada em SVMs, Lógica *Fuzzy* [Zad65] [Bro94], ou em outros algoritmos de AM. A combinação de classificadores tem como objetivo resolver problemas envolvendo diferentes tipos de dados, ou problemas não resolvidos satisfatoriamente por um único classificador e/ou em que o módulo de classificação não apresenta estabilidade. Os resultados de uma combinação de classificadores podem ser mais estáveis do que os fornecidos por um único classificador, pois a combinação de várias decisões é, geralmente, menos susceptível a oscilações. A combinação deve ressaltar as boas características de cada classificador e reprimir suas falhas [Sha97]. Uma consequência natural da exploração do potencial dos classificadores especialistas é considerar um sistema híbrido, no qual os módulos especialistas adotam diferentes técnicas para a solução do problema. A principal desvantagem da combinação de classificadores é o aumento da complexidade do problema, que pode ser tratada utilizando técnicas para particionar o problema entre os módulos da combinação [Ho94].

Uma possível configuração da arquitetura do módulo de análise com combinação de classificadores SVMs é apresentada na Figura 8.1. Na Figura 8.2 é sugerida uma outra configuração utilizando SVMs e redes MLP e RBF com as saídas sendo combinadas com Lógica *Fuzzy*. O objetivo, inicialmente, seria verificar se o desempenho do módulo de análise melhora com a combinação. Em caso afirmativo, diferentes arquiteturas e técnicas poderiam ser investigadas para a combinação.

Figura 8.1 – Possível arquitetura do módulo de análise utilizando combinação de classificadores.

Figura 8.2 – Sugestão de combinação de classificadores utilizando Lógica *Fuzzy*.

Referências Bibliográficas

- [Ana00] ANATOMIA OCULAR (2000). *Informações sobre a medicina dos olhos*. <http://www.oftalmo.com.br/anatomia.htm> (março).
- [App99] APPLGATE, R. & THIBOS, L. N. (1999). *Proposal for a VSIA taskforce on methods and standards for reporting aberration characteristics of eyes*. VSAI-99. Santa Fe, February, <http://www.opt.indiana.edu/people/faculty/thibos/VSIA/VSIA-99/sld001.htm>
- [Ash98] ASH, D. L. et al. (1998). Optimal hartmann sensing at low light levels, *Optics Communications*, vol. 156, pp. 10-15, November.
- [Bal98] BALDI, P. & BRUNAK, S. (1998). *Bioinformatics – the machine learning approach*. The MIT Press.
- [Bar96] BARNES, T. H. et al. (1996). High-resolution adaptive optics using an interference phase loop, *Optics Communications*, vol. 132, pp. 494-502, December.
- [Bar01] BARSKY, E. et al. (2001) RAYS (render as you see): vision-realistic rendering using hartmann-shack wavefront aberrations, Internal report, <http://www.cs.berkeley.edu/optical/vrender/papers/psf.pdf> (March).
- [Bis95] BISHOP, C. (1995). *Neural networks for pattern recognition*, Ed. Oxford Press.
- [Bit96] BITTENCOURT, G. (1996). *Inteligência Artificial: Ferramentas e Teoria*, Instituto de computação, Unicamp.
- [Bra98] BRAGA, A. P. et al. (1998). Fundamentos de redes neurais artificiais, *XI Escola Brasileira de Computação*.
- [Bra00] BRAGA, A. P., et al. (2000). *Redes neurais artificiais: Teoria e Aplicações*. Editora de Livros Técnicos e Científicos (LTC), Rio de Janeiro, 262 Pp.
- [Bre96] BREGA, J. R. F. (1996). *A utilização de redes neurais artificiais em um sistema de gerência de pavimentos*, Tese (Doutorado), EESC – USP, São Carlos, 234 p.
- [Bro94] BROWN, M. & HARRIS, C. (1994). *Neurofuzzy adaptive modelling and control*, Prentice Hall International.
- [Bro98] BROOMHEAD, D. S. & LOWE, D. (1988). *Multivariable functional interpolation and adaptative networks*, *Complex Systems*, 2:321-355.
- [Byu02] BYUN, H. & LEE, S. (2002). Applications of support vector machines for pattern recognition: a survey. *Proceedings lecture notes in computer science*, v. 2388, Springer-Verlag, pp. 213-236.
- [Can97] CANON (1997). *Specifications of Auto ref R-50 and RK-5*, Canon Inc. Medical Equipment Group, printed in Japan.
- [Can99] CANON (1999). *RK-5 auto ref-keratometer - Take the next step*, Canon Inc.,

<http://www.canom.com> (novembro).

- [Car98] CARVALHO, L. A. V. et al. (1998). Desenvolvimento de um instrumento computadorizado para medida do poder refrativo da córnea (videoceratógrafo), *Arquivo Brasileiro de Oftalmologia*, 61(6), dezembro.
- [Car99] CARVALHO, A. C. P. L. F. (1999). *Notas de aula: SCE5809 - Redes neurais*, Grupo de Inteligência Computacional, Pós-graduação em Computação, ICMC-USP, São Carlos-SP.
- [Car02] CARVALHO, B. P. R. (2002). *Support vector machines – um estudo sobre técnicas de treinamento*, Monografia, UFMG. <http://www.litc.cpdee.ufmg.br>
- [Cas03] CASTANÓN, C. A. B. (2003). Recuperação de imagens por conteúdo através de análise por wavelets, Dissertação (Mestrado), ICMC/USP, São Carlos, SP, 112 Pp, Fevereiro.
- [Cel99] CELLI, P. L. F. (1999). Um sistema de apoio à identificação de suspeitos com reconhecimento automático de faces. Dissertação (Mestrado), ICMC-USP, 100 Pp, São Carlos, SP.
- [Cen01] CENTRE FOR EYE RESEARCH (2001), *Ophthalmic and visual optics*, <http://www.hlth.qut.edu.au/opt/research/atchisonresearch.html> (June).
- [Cha76] CHARMAN, W. N. (1976). A pioneering instrument: the collin's electronic refraction meter, *Ophthalmic Optician*, v. 16, p.345.
- [Cha93] CHANG, T. & KUO, C. J. (1993). Texture Analysis and Classification with Tree-Structured - Wavelet Transform. *IEEE Transaction on Image Processing*, 2(4):429–441.
- [Che98] CHERKASSKY, V. & MULIER, F. (1998). *Learning from data. Concepts, theory and methods*. John Wiley and Sons, Inc.
- [Col02a] COLLOBERT, R. (2002). *A small user guide for svmtorch*. <http://www.idiap.ch/learning/SVMTorch.html> (agosto).
- [Col02b] COLLOBERT, R. & BENGIO, S. (2002). *Support vector machines for large-scale regression problems*. Research Report IDIAP-RR-00-17. <http://www.idiap.ch/learning/SVMTorch.html> (agosto).
- [Cos96] COSTA, J. A. F. (1996). *Sistema de reconhecimento de padrões visuais invariantes e transformações geométricas utilizando redes neurais artificiais de múltiplas camadas*. Dissertação (Mestrado), Escola de Engenharia de São Carlos - USP, São Carlos - SP, 157 Pp., janeiro.
- [Cox81] COX, J. (1981). Evaluation of peak location algorithms with subpixel accuracy for mosaic focal planes, in *Processing of images and data from Optical sensors*, W. H. Carter, ed., Proc. Soc. Photo-Opt. Instrum. Eng., 292, pp. 288-299.
- [Cri01] CRISTIANINI, N & SHAW-TAYLOR, J. (2001). *An introduction to support vector machines and other kernel-based learning methods*. Cambridge University Press.
- [Cub79] CUBALCHINI, J. (1979). Modal wave-front estimation from phase derivative measurements, *J. Opt. Soc. Am.* 69, 972-977.
- [Dau88] DAUGMAN, D. (1988). Complete Discrete 2-D Gabor Transforms by Neural Networks for Image Analysis and Compression, *IEEE Trans. on Acoustic, Speech, and Signal Processing*, 36(7): 1169–1179.
- [Dau90] DAUBECHIES, I. (1990). The Wavelet Transform, Time Frequency Localization, and Signal Analysis. *IEEE Transactions on Information Theory*, 36(5): 961–1005.

- [Dau92] DAUBECHIES, I. (1992). Ten Lectures on Wavelets. *CBMS-NSF Regional Conference Series in Applied Mathematics, Rutgers University and AT&T Bell Laboratories*, v. 61.
- [Dau95] DAUGMAN, J. G. (1995). Uncertainty Relation for Resolution in Space, Spatial Frequency and Orientation Optimized by Two-Dimensional Visual Cortical Filters. *Journal of Optical Society of America*, v. 2, No. 7, pp. 1160-1169.
- [Dau01] DAUGMAN, J. (2001). Gabor wavelets and statistical pattern recognition. *The handbook of Brian theory and neural networks*, 2nd ed., MIT Press (M. Arbib), pp 457-463.
- [Dau02] DAUGMAN, J. (2002). How Iris Recognition Works. *Proceedings of 2002 International Conference on Image Processing*, 1:33–36.
- [Day00] DAYTON, D. et al. (2000). Laboratory and field demonstration of a low cost membrane mirror adaptive optics system, *Optics Communications*, vol. 176, pp. 339-345, April.
- [Doe99] DOENÇAS OCULARES (1999). *Erros de refração – Correção visual a laser* <http://www.visuallaser.com.br/doencas.html> (novembro).
- [Dum98] DUMMAIS, S. (1998). Using SVM for text categorization. *IEEE Intelligence Systems*, pp. 21-23.
- [Egm02] EGMONT-PETERSEN, M. et al. (2002) Image processing with neural networks – a review, *Pattern Recognition*, v. 35, pp. 2279-2301, October.
- [Err99] ERROS DE REFRAÇÃO (1999). *O que é um erro de refração?*, <http://www.ggvc.com.br/laserocular/errosde.htm> (novembro).
- [Fac01] FACELI, K. (2001). *Combinação de métodos de inteligência artificial para fusão de sensores*. Dissertação (Mestrado), ICMC - USP, São Carlos - SP, 156 Pp., fevereiro.
- [Fah88] FAHLMAN, S. E. (1988). *Faster-learning variations on back-propagation: An empirical study*. Connectionist Models Summer School.
- [Fas01] FAST RETINA SCANNER (2001). <http://wwwasic.kip.uni-heidelberg.de/wavefront/> (June).
- [Fer99] FERNADEZ, R. & VIENNET, E. (1999). Face identification using support vector machines, *Proceedings of the European symposium on artificial neural networks*, pages 195-200, Brussels, D-Facto Press.
- [Fer01] FERIS, R. S., et al. (2001). Efficient Real-Time Face Tracking in Wavelet Subspace. *Proceedings of International Conference on Computer Vision, Workshop on Recognition, Analysis and Tracking of Faces and Gestures in Real-Time Systems, Vancouver, BC.*, pp. 113–118.
- [Fit99] FITZPATRICK, J. M. & SONKA, M. (1999). *Handbook of Medical Imaging: volume 3 process in medical image processing and analysis*, SPIE Optical Engineering Press, September.
- [For98] FORESTI, G. L. & PIERONI, G. (1998). Exploiting neural trees in range image understanding, *Pattern Recognition Letters*, 19 (9), pp. 869-878.
- [Fun01] FUNDINGSLAND, B. (2001). *The developing science of wavefront ablations*. <http://www.eyeworld.org/apr01/0401p59.html> (April).
- [Gab46] GABOR, D. (1946). Theory of Communication. *Journal of the Institute of Electrical Engineers*, v. 93, pp. 429-457.

- [Gal00] GALO, M. L. B. T. (2000). *Aplicação de redes neurais artificiais e sensoriamento remoto na caracterização ambiental do parque estadual morro do diabo*, Tese (Doutorado), Escola de Engenharia de São Carlos – USP, São Carlos, 204 p.
- [Glo99] GLOSSARY (1999). *Glossary about schematic eye – webvison master* – <http://spectable.berkeley.edu/cleere/glossary> (novembro).
- [Gob99] GOBBI, P. G., et al. (1999). *PRK ablation algorithms: customized or default corneal curvature?*, Laser medicine research and department of ophthalmology and visual sciences, Scientific Institute San Raffaele Hospital, University of Milan-Italy, *Annual Meeting*, Fort Lauderdale, Florida - may 9-14.
- [Gon92] GONZALEZ, R. C. & WOOD, R. E. (1992). *Digital image processing*, Reading, Ed. Addison-Wesley.
- [Gua94] GUARDO, R & ADLER, A. (1994). A neural network image reconstruction technique for electrical impedance tomography, *IEEE Trans. Medical Imaging*, 13 (4), pp. 594 - 600.
- [Gun98] GUNN S. (1998). Support Vector Machines for classification and regression, ISIS technical report, University of Southampton, 52 Pp.
- [Guy81] GUYTON, D. L. (1981). Automated Clinical Refraction. In: DUANE, T. D. *Clinical Ophthalmology*, Harper & Row Publishers, chap. 67, pp. 1.
- [Ham97] HAMAM, H. & SIMONET, P. (1997). *Measurement of the human eye aberrations using FLC device*, FLC 97. http://www-opt.enst-bretagne.fr/hamam/publi_abstracts.html (June).
- [Ham00] HAMAM, H. (2000). A direct technique for calculating the profile of aberration of the eye measured by a modified hartmann-shack apparatus, *Vision Research*, vol. 39, pp. 4309-4323.
- [Ham01] HAMAM, H. (2001). A quick method form analyzing Hartmann-Shack patterns: application to refractive surgery, *Journal of refractive surgery*, <http://www.slackinc.com/eye/jrs/vol165/hamam.htm> (June)
- [Hay94] HAYKIN, S. (1994). *Neural Networks – A comprehensive foundation*, Ed. IEEE Press, 592 Pp.
- [Har78] HARDY, J. W. (1978). Active optics: a new technology for the control of light, *Proceeding IEEE*, (66) 651-697.
- [Hec89] HECHT-NIELSEN, R. (1989). Theory of the Back-propagation neural network, *Proceeding IJCNN*, Washington D.C., pp. 18-22.
- [Her91] HERTZ, J. et al. (1991). *Introduction to the theory of neural computation*, vol. 1, Santa Fé Institute studies in the sciences of complexity, Ed. Addison-Wesley Publishing Company.
- [Him01] HIMEBAUGH, N. et al. (2001). *Comparison of Fluorescein and Hartmann-Shack wavefront sensing methods for monitoring tear film breakup*, <http://research.opt.indiana.edu/>
- [Ho94] HO, T. K. et al. (1994). Decision combination in multiple classifier systems, *IEEE Transaction on pattern analysis machine intelligence*, v. 16, n.1, p. 66-75.
- [Hol75] HOLLAND, J. H. (1975). *Adaptation in natural and artificial systems*. MIT Press.
- [Hon01] HONG, X. (2001). *Optical aberrations of human eyes and their impact on visual performance*, PhD Dissertation, School of Optometry, Indiana University, 335Pp., August.

- [Hon01a] HONG, X. & THIBOS, L. N. (2001). *Optical aberrations following laser in situ keratomileusis (LASIK) surgery*, <http://research.opt.indiana.edu/Library> (January).
- [Hon02] HONG, X. et al. (2002). *Shack-Hartmann data analysis software for MATLAB* <http://research.opt.indiana.edu/Library/SHtoolkit/README.html>
- [Hum98] HUMPHREY (1998). *Humphrey refractor/keratometer hark 599 specifications*, <http://www.humphrey.com/products/hark599.htm> (agosto).
- [Jai01] JAIN, A., et al. (2001). Fingerprint Matching using Minutiae and Texture Features. *Technical Report MSU-CSE-01-17*, Department of Computer Science, Michigan State University, East Lansing, Michigan.
- [Joa99] JOACHIMS, T. (1999). Making large-scale SVM learning practical. In: Schölkopf et al., *Advances in kernel methods – support vector learning*. MIT Press
- [Jon82] JOHNSON, R.A. & WICHERN, D. (1982). *Applied multivariate statistical analysis*, 2nd edition, Prentice-Hall, New Jersey.
- [Ker02] KERNEL (2002). *Kernel machines*, <http://kernel-machines.org> (February).
- [Kin97] KING, D., CAMPBELL, M., HAMAM, H. (1997). *Aberrations of the human eye: comparative objective and subjective measurements*, OSA97. http://www-opt.enst-bretagne.fr/hamam/publi_abstracts.html (June).
- [Kle98] KLEIN, S. A. (1998). Optimal corneal ablation for eyes with arbitrary Hartmann-Shack aberrations, *J. Opt. Soc. Am.*, vol. 15, no. 9/ September, pp. 2580-2588.
- [Lac98] LACERDA, E. G. M. (1998). *Otimização de redes neurais RBF usando algoritmos genéticos e sua aplicação na área financeira*. Dissertação (Mestrado), Instituto de Ciências Matemáticas e de Computação - USP, São Carlos - SP, 158 Pp., Dezembro.
- [Lac01] LACERDA, E. et al. (2001) Evolutionary optimization of RBF Networks, *International Journal of Neural Systems*, vol. 11, n. 3, pp. 287-294, World Scientific Pub Co.
- [Law96] LAW, N. F. & LANE, R. G. (1996). Wavefront estimation at low light levels, *Optics Communications*, vol. 126, pp. 19-24, may.
- [Lec95] LECUN, Y. A et al. (1995). Comparison of learning algorithms for handwritten digit recognition. *Proceedings of the international conference on artificial neural networks*, vol. 2, pages 53-60, Nanterre, France.
- [Lee90] LEE, J. et al. (1990). A neural network approach to cloud classification, *IEEE Transaction on geosciences and remote sensing*, V. 28, n. 5, pp. 846-855.
- [Lee96] LEE, T. S. (1996). Image Representation Using 2D Gabor Wavelets. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(10):959–971.
- [Lee02] LEE, S. & VERRI, A. (2002). Pattern recongition with support vector machines, First Internacional Workshop - SVM02, Niagara Falls, Canada, August., Proceedings Lecture Notes in Computer Science, n. 2388, Springer-Verlag.
- [Leo94] LEOW, W. K. (1994). *Visor: learning visual schemas in neural networks for object recognition and scene analysis*, thesis (doctorate), University of Texas, Austin.
- [Ler98] LERNER, B. (1998). Toward a completely automatic neural network based human chromosome analysis, *IEEE Trans. System, Man, Cybernet. Part B: Cybernetics*, 28 (4), pp. 544-552.

- [Lia91] LIANG, J., GRIMM, B., GOELZ, S., BILLE, J. F., (1991). Hartmann-Shack sensor as a component in an active optical system to improve the depth resolution of the laser tomographic scanner, *Active and Adaptive Optical Systems*, M. A. Ealey, ed., Proc. Soc. Photo-Opt. Instrument Eng. - SPIE, vol. 1542, pp. 543-554.
- [Lia94] LIANG, J., GRIMM, B., GOELZ, S., BILLE, J. F., (1994). Objective measurement of wave aberrations of the human eye with the use of a Hartmann-Shack wave-front sensor. *J. Opt. Soc. Am.*, Vol. 14, No. 11/ July, pp. 1949-1957.
- [Lia97a] LIANG, J., WILLIAMS, D. R., (1997). Aberrations and retinal image quality of the normal human eye. *J. Opt. Soc. Am.*, Vol. 14, No. 11/ November, 2873-2883.
- [Lia97b] LIANG, J., WILLIAMS, D. R., MILLER, D. T., (1997). Supernormal vision and high-resolution retinal imaging through adaptive optics, *J. Opt. Soc. Am.*, Vol. 14, No. 11/ November, 2884-2892.
- [Lip87] LIPPMANN, R. P. (1987). An introduction to computing with neural nets, *IEEE ASSP Magazine*, v. 4, pp. 4-22.
- [Löf96] LÖFDAHL, M. G. (1996). *Phase diversity wavefront sensing and image restoration applied to high-resolution solar observations*, Doctoral dissertation, Stockholm observatory, University of stockholm. <http://www.astro.su.se/~mats/publications/dissertation.html>
- [Lor02] LORENA, A. C. & CARVALHO, A. C. P. L. F. (2002). As máquinas de vetores suporte, relatório interno, Laboratório de Inteligência Computacional, ICMC/USP, São Carlos – SP, junho.
- [Lus93] LU, S. & SZETO, A. (1993). Hierarchical artificial neural networks for edge enhancement. *Pattern Recognition*, vol. 26, n. 8.
- [Mal99] MALLAT, S. G. (1999). *A Wavelet Tour of Signal Processing*. Academic Press, 2nd. Edition.
- [Man86] MANLY, B. F. J. (1986). Multivariate statistical methods, Chapman and hall, London.
- [Man96] MANJUNATH, B. & MA, W. (1996). Texture Features for Browsin and Retrieval of Image Data, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 18:837–842.
- [Mar92] MARTINS, F. C. M. (1992). *Aplicação de redes neurais em processamento de imagens. Detecção de bordas através de cooperação e competição*, Dissertação (Mestrado), Escola Politécnica, São Paulo - SP, 153 Pp, junho.
- [Mas89] MASON, R., GUNST, R., e HESS, J. (1989). *Statistical design and analysis of experiments*, John Wiley & Sons, 330 Pp.
- [Mic00] MICHAELIS (2000). Moderno Dicionário da Língua Portuguesa e antônimos. <http://www.uol.com.br/bibliot/dicionar/> (novembro).
- [Mil96] MILLER, D. T. et al. (1996). Images of cone Photoreceptors in the living human eye, *Vision Res.*, vol. 36, n.8, pp. 1067-1079.
- [Mit97] MITCHELL, T. (1997). *Machine Learning*. McGraw Hill.
- [Moh95] MOH, J. & SHIH, F. (1995). A general-purpose model for image operations based on multilayer perceptrons, *Pattern Recognition*, vol. 27, n. 3.
- [Mor91] MORIMOTO, C. H. (1991). *Sistema de visão robótica para o reconhecimento de objetos*

poliédricos a partir de uma única imagem. Dissertação (Mestrado), Escola Politécnica - USP, São Paulo - SP, 156 Pp., agosto.

- [Mou97] MOURA JR., L. A. (1997). Processamento de imagens médicas – visão geral: introdução, *Fenasoft'97 - Congresso aberto para profissionais da área de saúde*, Escola Politécnica – USP, julho.
- [Nak02] NAKAJIMA, C. et al. (2002). Full-body person recognition system, *Pattern Recognition*, September, 15 Pp.
- [Ned02] NEURODIMENSION (2002). *NeuroSolutions*. <http://www.nd.com/> (September).
- [Nid98] NIDEK (1998). *Auto refractometers AR-600, ARK-700 and AR-660*, <http://www.nidek.com>, (agosto).
- [Nik98] NIKON (1998). *NR5500 Auto Refractometer - Ophthalmic instruments*, <http://www.nikonusa.com/products/products.taf?id=103> (agosto).
- [Oko00] OKO TECHNOLOGIES (2000). *Microlens arrays for Hartmann-Shack sensors*. <http://www.okoteck.com/>
- [Olh99] OLHO (1999). *Informações gerais sobre o olho humano*, Laser Ocular Brasil-Canadá. <http://www.ggvc.com.br/laserocular/olho.htm> (novembro).
- [Olh00] OLHO (2000). *Informações sobre as partes que constituem o olho humano*, <http://rot.fmrp.usp.br/bo/paginas/olho.htm> (abril).
- [Oli97] OLIVEIRA, P. R. (1997). *Redes neurais artificiais para extração de características*, Dissertação (Mestrado), ICMC - USP, São Carlos – SP, 98 Pp.
- [Opt02] OPTICAL (2002). *Optics and Topography involving the cornea and lens*. <http://www.cs.berkeley.edu/optical> (December).
- [Osu98] OSUNA, E. (1998). Applying SVMs to face detection, *IEEE Intelligence Systems*, pp. 23-26.
- [Pal93] PAL, N. R & PAL, S. K. (1993). A review on image segmentation techniques, *Pattern Recognition*, 26 (9), pp. 1277-1294.
- [Pas98] PASQUARIELLO, G. (1998). Automatic target recognition for naval traffic control using neural networks, *Image Vision Computer*, 16 (2), pp. 67-73.
- [Per94] PEREZ, L. G. et al. (1994). Training an artificial neural network to discriminate between magnetizing inrush and internal faults, *IEEE Transactions on Power Delivery*, vol. 9, n. 1, pp. 434-441.
- [Per95] PERELMUTER, G. et al. (1995). Reconhecimento de imagens bidimensionais utilizando redes neurais artificiais, *Anais do VIII SIBGRAPI*, pp. 197-203.
- [Pet99] PETTIT, G. H., et al. (1999). *Customized corneal ablation: wavefront guided laser vision correction*, AUTONOMOUS TECHNOLOGY (ORLANDO, FL) ARVO, *Annual Meeting*, Fort Lauderdale, Florida (USA) – may, 9-14.
- [Rei94] REINUS, W. R. et al. (1994). Diagnosis of focal bone lesions using neural networks, *Invest. Radiol*, 29 (6), pp.606-611.
- [Rie93] RIEDMILLER, M. & BRAUN, H. (1993). A direct adaptative method for faster Back-propagation learning: The RPROP algorithm, *Proceedings of the IEEE International Conference on Neural Networks (ICNN 93)*.

- [Rio97] RÍOS, S. et al. (1997). Hartmann sensing with albrecht grids, *Optics Communications*, v. 133, pp. 443-453.
- [Rip96] RIPLEY, B. D (1996). *Pattern Recognition and neural networks*, Cambridge University Press.
- [Rom97] ROMDHAMI, S. (1997). *Face recognition using principal components analysis*, University of Glasgow.
- [Roo97] ROOIJ, A. J. F. et al. (1997). *Neural network training using genetic algorithms*, Series in Machine Perception and Artificial Intelligence, vol. 26, World Scientific Pub Co.
- [Rum86] RUMELHART, D.E et al. (1986). Learning internal representations by error propagation. In: D.E Rumelhart e J.L McClelland. Eds., *Parallel distributed processing: explorations in the microstructure of cognition*, v.1, Cambridge, MA. MIT Press.
- [Rup02] RUPING, S. (2002). *Mysvm-manual*. <http://www-ai.cs.uni-dortmund.de/SOFTWARE/MYSVM/> (agosto).
- [Saf81] SAFIR, A. (1981). Retinoscopy In: DUANE, T. D. *Clinical Ophthalmology*, Philadelphia, Harper & Row Publishers. Revised Edition, Chap. 37, p.1-17.
- [Sal81] SALOMON, P. M. (1981). Charge-coupled device trackers for high-accuracy guidance applications, *Optical Engineering*, January/February, vol. 20, n. 1, pp. 135-142.
- [Sal99] SALAMA, N. H. et al. (1999). Wavefront sensor using the talbot effect, *Optics & laser technology*, vol. 31, pp. 269-272.
- [Sam01] SAMALONIS, L. (2001). *The new refractive technologies*. <http://www.eyeworld.org/apr01/0401p68.html> (April)
- [San02] SANTOS, E. M. & GOMES, H. M. (2002). A comparative study of polynomial kernel SVM applied to appearance-based object recognition, *Proceedings lecture notes in computer science*, v. 2388, Springer-Verlag, pp. 408-418.
- [San98] SANTOS, J. F. et al.. (1998). Estudo comparativo entre dois métodos de refração: objetiva e subjetiva, *Revista Brasileira de Oftalmologia*, vol. 57, n. 2, fevereiro.
- [Sau98] SAUNDERS, C., et al. (1998). *Support vector machine - reference manual*. Technical Report CSD-TR-98-03, Royal Holloway, University of London. <http://svm.dcs.rhnc.ac.uk/>
- [Sch96] SCHOLKOPF, B., et al. (1996). *Comparing support vector machines with gaussian kernels toradial basis functions classifier*. <http://www.kernel-machine.org/publications.html> (august, 2000).
- [Sch98] SCHÖLKOPF, B. (1998). SVMs – a practical consequence of learning theory. *IEEE Intelligence Systems*, pp. 36-40, July.
- [Ser88] SERRA, J. P. F. (1988). *Image analysis and mathematical morphology*, volume 2: theoretical advances. Academic Press, London, 411 Pp.
- [Sha71] SHACK & PLATT (1971). Production and use of a lenticular Hartmann screen, *JOSA*, no. 61, pp. 656-670.
- [Sha97] SHARKEY, A. J. C. (1997). Modularity, combination and artificial neural nets, *Connection Science*, v. 9, n.1, pp. 3-10.

- [Sim98] SIMÕES, M. G. (1998). *Automação de processos baseada em controle e modelagem inteligente*. Tese (Livre Docência), EP - USP, São Paulo - SP, 218 Pp., Junho.
- [Slo70] SLOANE, A. E. (1970). *Manual of refraction*, Boston, Little & Brown.
- [Sno99] SNOWDEN, S. (1999). *OD's are solving problems and tooling up for the new millennium with cutting-edge equipment for their exam and labs*, Feature Story, <http://www.optcom3.com/sojo/sojo1/feature.html> (November).
- [Son99] SONKA, M.; HLAVAC, V.; BOYLE, R. (1999). *Image processing, analysis, and machine vision*. Second ed. PWS Publishing.
- [Sou80] SOUTHWELL, W. H. (1980). Wavefront estimation from wavefront slope measurements, *Journal Opt. Soc. Am*, 70, pp. 998-1006.
- [Sta96] STASSOPOULOU, A. et al. (1996). Bayesian and neural networks for geographic information processing, *Pattern Recognition Lett.*, 17 (13), pp. 1325-1330.
- [Ste99] STEINBERG, C. S. (1999). *Eye exam tests*. <http://www.visioncare.com/cityeyes/tests.htm> (November).
- [Tat00] TATIBANA, C. Y. & KAETSU, D. Y. (2000). Uma introdução às redes neurais. <http://www.din.uem.br/ia/neurais/> (março).
- [Tec99] TECNOLOGIAS EM TRATAMENTO (1999). *Tipos de tratamento para correção de vícios oculares*. <http://www.ggvc.com.br/laserocular/novidade.htm> (novembro).
- [Thi99a] THIBOS, L. N. (1999) *Perfecting the optical system of the human eye*, IU Research Lecture, School of Optometry, Indiana University, Bloomington, April. http://www.opt.indiana.edu/people/faculty/thibos/LNT_DFL/slide.html
- [Thi99b] THIBOS, L. N. & XIN HONG (1999). Clinical applications of the shack-hartmann aberrometer, *Optometry and vision science*, vol. 76, no. 12, pp. 817-825.
- [Thi00a] THIBOS, L. N. (2000) *Principles of hartmann-shack aberrometry*, Wavefront Sensing Congress, Santa Fe. http://www.opt.indiana.edu/people/faculty/thibos/VSIA/VSIA-2000_SH_tutorial_v2/index.htm
- [Thi00b] THIBOS, L. N. et al. (2000) *Standards for reporting the optical aberrations of eyes*, OSA Taskforce Members, VSIA-2000, <http://www.osa.org/homes/vision/resources/%7Ehsbiblio.htm> (May).
- [Thi02a] THIBOS, L. N. (2002). *Research Computing Laboratory – School of Optometry*, <http://research.opt.indiana.edu/>
- [Thi02b] THIBOS, L. N., et al. (2002). Statistical variation of aberration structure and image quality in a normal population of healthy eyes, *Journal of the Optical Society of America A*, Volume 19, Issue 12, pp. 2329-2348.
- [Tin99] TINÓS, R. (1999). *Detecção e diagnóstico de falhas em robôs manipuladores via redes neurais artificiais*, Dissertação (Mestrado), EESC – USP, São Carlos, 117 Pp.
- [Top98] TOPCON (1998). *Bringing the future into focus*, <http://www.topcon.com/auto.htm> (August).
- [Ton02] TONICE, A. (2000). *Processamento automáticos de ações faciais superiores usando redes neurais artificiais e processamento de imagens*. Dissertação (Mestrado), Escola Politécnica - USP, São Paulo - SP, 110 Pp., março.

- [Tor02] TORCH (2002). *Torch: a modular machine learning software library*. <http://www.torch.ch/>
- [Ura95] URAS, R. (1995). *Análise comparativa da refração automática objetiva com refração clínica*, Tese (Doutorado), UNIFESP, Escola Paulista de Medicina, São Paulo – SP, 79 Pp.
- [Val01] VALERIO NETTO, A. (2001). *Proposal to use artificial intelligence in the field of ophthalmology*, Oxyopia: Vision Science Seminar – Indiana University (IU), Indiana - USA, October.
- [Val02] VALERIO NETTO, A. & OLIVEIRA, M. C. F. (2002). Using neural nets to measure ocular refractive errors - a proposal, *Proceeding SPIE - Applications and Science of Neural Networks, Fuzzy Systems, and Evolutionary Computation*, Seattle – EUA, December.
- [Vas95] VASCONCELOS, G. C. (1995). Redes neurais e reconhecimento de padrões. *Tutorial 2*. In: II Simpósio Brasileiro de Redes neurais, São Carlos - SP.
- [Vap68] VAPNIK, V. N. & CHERVONENKIS, A. (1968). On the uniform convergence of relative frequencies of events to their probabilities. *Theory of probability and applications*, (16):262-280.
- [Vap92] VAPNIK, V. et al. (1992). A Training Algorithm for Optimal Margin Classifiers. In *Fifth Annual Workshop on Computational Learning Theory*, Pittsburgh.
- [Vap95] VAPNIK, V. (1995). *The Nature of Statistical Learning Theory*, Springer-Verlag, New York.
- [Ven93] VENTURA, L. et al., (1993). Desenvolvimento de um sistema de detecção de luz I.V. espalhada pela retina, *Rev Fís Aplicada a Instrumentação*, v. 8, n. 3, p. 111-116, setembro.
- [Ven94a] VENTURA, L. et al., (1994). Development of an automatic system for ocular refraction measurements. *Phys. in Medicine and Biology* 39a, pp. 146.
- [Ven94b] VENTURA, L. et al., (1994). Desenvolvimento de um sistema para diagnóstico de erros refrativos oculares, *Rev. Fís. Aplicada Instrumental*, v. 9, n. 2, pp. 55-60, Junho.
- [Ven95] VENTURA, L. (1995). *Sistema de medidas automáticas de vícios de refração ocular*, Tese (Doutorado), Instituto de Física de São Carlos - USP, São Carlos - SP, 104 Pp., Outubro.
- [Ven96] VENTURA, L. et al. (1996). Desenvolvimento de um *software* para a determinação automática de miopia, hipermetropia e astigmatismo oculares, *Revista de Física Aplicada e Instrumentação*, vol. 11, n. 2, pp. 77-83, junho.
- [Vis01] VISUAL OPTICS LAB (2001). <http://research.opt.indiana.edu/Labs/427/427.html>
- [Wal84] WALSH, G. et al. (1984). Objective technique for the determination of monochromatic aberrations of the human eye, *Journal of Optical Society of America*, vol. 1, pp. 987-992.
- [Wav00] WAVEFRONT SCIENCES OPTICS (2000). *The shack-hartmann sensor*, Albuquerque, New Mexico. <http://wavefrontsciences.com> (May).
- [Wav01] WAVESCAN (2001). *Raising the standard of vision to new heights*, VISX Wavefront, <http://www.opto.com.br/english/medicaldiv/products/WaveScan.htm> (March).
- [Wav02a] WAVELET (2002) <http://www.barrt.ru/parshukov/about.htm> (December).
- [Wav02b] WAVELET COMPRESSION (2002) http://www.bearcave.com/misl/misl_tech/wavelets/compression/ (December).

- [Wav02c] WAVELET (2002) <http://www.seyboldreports.com/SRIP/wavelet/> (December).
- [Web92] WEBB, R. et al. (1992). Measurement of ocular local wavefront distortion with a spatially resolved refractometer, *Application. Opt.*, vol. 31, pp. 3678-86.
- [Wei98] WEISS, S. M. & N. INDURKHYA (1998). *Predictive Data Mining: A Practical Guide*, Morgan Kaufmann Publishers, Inc., San Francisco, CA.
- [Wel96] WELL, G. et al. (1996). Promising research: vision-based robot positioning neural networks, *Image Vision Computer*, 14 (10), pp. 715-732.
- [Wil92] WILLIAMS, H. P. et al. (1992). *Numerical recipes in C – the art of scientific computing*, second edition, Cambridge University Press, 979 Pp.
- [Wil96] WILLIAMS, D. R. et al. (1996). Off-axis optical quality and retinal sampling in the human eye, *Vision Res.*, vol. 36, n.8, pp. 1103-1114.
- [Woo96] WOODS, R. L. et al. (1996). Monocular diplopia caused by ocular aberrations and hyperopic defocuses, *Vision Res.*, vol. 36, no. 22, pp. 3597-3606.
- [Zad65] ZADEH, L. A. (1965). *Fuzzy sets*, *Information and control*, vol. 8, pp. 338-353.
- [Zel95] ZELL, A et al. (1995). *SNNS: Stuttgart Neural Network Simulator*. User manual, version 4.1, University of Stuttgart. (Report No. 6/95).

Bibliografia Consultada

- ADELI, H. HING, S. (1995). *Machine learning: neural network, genetic algorithms and fuzzy systems*, Ed. John Wiley & Sons.
- ALEKSANDER, I. & MAMDANI, E. H. (1968). Microcircuit learning nets: improved recognition by means or pattern feedback. *Electronics Letters*, 4(20), pp. 425-426.
- AVILA, M. et al. (1997). Integração de dois sistemas digitais de angiografia e microperimetria da retina e coróide, *Revista Brasileira de Oftalmologia*, vol. 56, n.2, fevereiro.
- AZEVEDO, F. M., et al. (2000). *Redes neurais com aplicações em controle e em sistemas especialistas*, Ed. Visual Books, 401Pp., Agosto.
- BARBOSA, A. V. (2001). *Análise de Componente Principal*, <http://www.cpdee.ufmg.br/~vilela/visão/node3.html> (novembro).
- BARBOZA, W. L. & SUSANNA Jr., R. (1998). Estudo de valores estereométricos em olhos graucomatosos e normais pela tomografia neuro-retiniana, *Revista Brasileira de Oftalmologia*, vol. 57, n. 3, março.
- BARNETT, J. A. (1981). *Computational methods for a mathematical theory of evidence*, IFCAI paper.
- BARTLETT, M. S., et al. (1996). *Classifying facial actions advances*, IN: Neural Information Processing Systems, MIT Press, Cambridge, MA, pp. 823-829.
- BARTLETT, M. S., et al. (1999). *Measuring facial expressions by computer image analysis psychophysiology*, 36:(2), pp. 253-263, March.
- BEUCHER, S. & MEYER, F. (1993). The morphological approach to segmentation: the watershed transformation. In: *Mathematical Morphology in Image Processing*, pp. 433-481, Marcel Dekker, New York.
- BOLLE, R. M. & VEMURI, B. C. (1991). On three-dimensional surface reconstruction methods, *Applications on pattern analysis and machine intelligence*, vol. 13, n. 1, January, pp. 1-13.
- BOOKSTEIN, F. L. (1997). Thin-plate splines and the atlas problem for biomedical images, *Lecture notes in computer science*, Information processing in medical imaging.
- BRADLEY, A. & THIBOS, L. (2000). *How to measure chromatic aberration and locate useful reference axes of the human eye*, OSA'95 Conference. <http://www.opt.indiana.edu/people/faculty/thibos/ABLNTOSA95/slide01.html> (June).
- BRODLIE, K. W. et al. (1992). *Scientific visualization – techniques and applications*, Springer-Verlag, 284 Pp.
- BROWN, C. W. & SHEPHERD, B. J. (1995). *Graphics file formats – reference and guide*, Ed. Prentice Hall, 472 Pp.

- CAMPBELL, M. C. et al. (1990). Psychophysical measurement of the blur on the retina due to optical aberrations of the eye, *Vision Research*, vol. 30, pp. 1587-1602.
- CARVALHO, L. A. V. (1996). *Desenvolvimento de um instrumento computadorizado para medida da curvatura da córnea durante o ato cirúrgico*. Dissertação (Mestrado), Instituto de Física de São Carlos - USP, São Carlos - SP, 115 Pp.
- CHEN, Y. et al. (2001). One-Class SVM for learning in image retrieval, *IEEE International Conference on Image Processing*, Thessaloniki, Greece, pp.54-58.
- COHEN, P. R. (1995). *Empirical methods for artificial intelligence*, Mit Press, Cambridge, MA, USA.
- COIERA, E. W. (1999). Inteligência Artificial na Medicina, *Informática Médica*, vol. 1, n.4, julho/agosto 1998, <http://www.epub.org.br/informaticamedica/n0104/coira.html> (outubro).
- COSTA, J. B. M. et al. (2002). *Highly accurate three-dimensional measurement of the optical and biomechanical properties of the human eye*. http://www.cs.berkeley.edu/~barsky/vis_sci_pubs.html (January).
- EATON, J. W. (1997). GNU Octave – a high-level interactive language for numerical computations, Edition 3, Free Software Foundation Editor, February.
- FRANÇA, C. A. (1999). *Avaliação da qualidade de placas de madeira através de um sistema de interferência nebulosa baseado em redes adaptativas*, Tese (Doutorado), Instituto de Física de São Carlos - USP, São Carlos - SP, 135 Pp., agosto.
- FUKUSHIMA, K. (1975). Cognitron: a self-organizing multilayered neural network. *Biological Cybernetics*, vol. 20, pp. 121-136.
- FÜSSEL, D. et al. (1997). Closed-loop fault diagnosis based on a no-linear process model and automatic fuzzy rule generation, *Proceedings of the IFAC Symposium on fault detection, Supervision and Safety for Technical Processes*, Kingston Upon Hull, U. K. vol. 1, pp. 359-364.
- GONZALEZ, R. C. & WOOD, R. E. (2000). *Processamento de imagens digitais*, Edição revisada e traduzida pelos Profs. Roberto M. Cesar Jr. e Luciano F. Costa, Ed. Edgard Blucher Ltda.
- GROSSBERG, S. (1976). Adaptive pattern classification and universal recoding: I. parallel development and coding of neural feature detectors. *Biological Cybernetics*, 23, pp. 121-134.
- HAMAM, H. et al. (1997). *A more flexible hartmann-shack apparatus for the measurement of aberrations of the human eyes*, OSA97. http://www-opt.enst-bretagne.fr/hamam/publi_abstracts.html (June).
- HARALICK, R. M. et al. (1987). Image analysis using mathematical morphology, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 9, no. 4, July.
- HEBB, D. O. (1949). *The organization of Behavior*, Ed. Wiley.
- HO, T. K. et al. (1994). Decision combination in multiple classifier systems, *IEEE Trans. on Pattern Analysis Machine Intelligence*, vol. 16, n. 1, pp. 66-75.
- HOPFIELD, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities, *Proceedings of the National Academy of Sciences, USA*, April, pp. 2554-2558.
- HUANG, J.& WECHSLER, H. (1999). Eye detection using optimal wavelet packets and radial basis functions (RBFs), *Int. Journal of Pattern recognition and artificial Intelligence*, vol. 13, n. 7.

- KLINGER, A.; FU, K. S.; Kunii T. L. (1977). *Data structures, computer graphics, and pattern recognition*, Ed. Academic Press, New York, 498 Pp.
- KOENDERINK, J. J. (1984). The Structure of Images, *Biological Cybernetics*, n.50, pp. 363-370.
- KOH, J. et al. (1995). A multilayer self-Organizing feature map for range image segmentation. *Neural Networks*, vol. 8, n.1.
- KOHONEN, T. (1989). *Self-Organization and Associative Memory*, Springer Verlag.
- KNIGHT, J. C. & LEVESON, N. G. (1986). An experimental evaluation of independence in multiversion programming, *IEEE Transaction on software engineering*, SE-12.
- KUNG, S. Y. & TAUR J. S. (1995). Decision based neural networks with signal image classification applications, *IEEE Trans. Neural Networks*, 6 (1), pp. 170-181.
- LAND, J. E. (2000). *Shack-Hartmann wavefront sensor for tomographic reconstruction of aerodynamic flow density measurements*, Thesis (Doctorate), University of Florida, 194 Pp., December.
- LANZARINI, L. & DE GIUSTI, A. (2000). Pattern recognition in medical images using neural networks, *Journal of Computer Science & Technology*, vol.1, no. 4.
- LE, D. X. et al (1995). Classification of binary document images into textual or no textual data blocks using network models, *Machine Vision Application*, 8 (5), pp. 289-304.
- LE GRAND, E. & EL HAGE, S. G. (1980). *Physiological Optics*, Springer Series in Optical Sciences, Springer-Verlag, vol. 13, 335 Pp.
- LESTER, H. & ARRIDGE, S. R. (1999). A survey of hierarchical non-linear medical image registration, *Pattern Recognition*, no. 32, pp.129-149.
- LIM, J. S. (1990). *Two-dimensional signal and image processing*, Prentice-Hall, Englewood Cliffs, NJ.
- LINDERBERG, T. (1995). A scale selection principle for estimating image deformations, *Proceeding 5th International Conference on Computer Vision*, Cambridge, MA, pp. 134-141, June.
- LINHARES, O. L. (1987). Matrizes de teste inteiras com inversas inteiras e autovalores e autovetores inteiros facilmente explicáveis, *Notas técnicas do ICMC/USP*, no. 25, 41 Pp., Abril.
- LUPPE, M. (1997). *Desenvolvimento de um processador pipeline dedicado para extração de bordas em tempo real*, Dissertação (Mestrado), IFSC - USP, São Carlos - SP, 91 Pp., Junho.
- MARTINS, L. D. & APPEL, E. (1997). Análise da imagem digital oftalmológica, *Revista Brasileira de Oftalmologia*, vol. 56, n.9, setembro.
- MCBRIAN, N. A. & MILLODOT, M. (1986). The effect of refractive error on the accommodative response gradient, *Ophthalmic. Physiol. Opt.*, v. 6, n.2, pp. 145-149.
- MCCULLOCH, W. S. & PITTS, W. H. (1943). A logical calculus of the ideas in nervous activity, *Bull Math Biophys*, 5, pp. 115-133.
- MILANOVA, M. et al. (1999). Applications of cellular neural networks for shape from shading problem, *International workshop on machine learning and data mining in pattern recognition, Lecture notes in artificial intelligence*, Springer-Verlag, pp. 52-63, Leipzig, Germany, September.
- MINSKY, M. & PAPERT, S. (1969). *Perceptrons*. MIT Press, Cambridge.

- MOURA JR., L. A. et al. (1999). *Processamento de imagens médicas – Formação de imagens médicas, princípios físicos e modelos*, Instituto do coração (INCOR) <http://www.incor.usp.br/spdweb/course.html> (novembro).
- MURINO, V. (1998). Structured neural networks for pattern recognition, *IEEE Trans. on systems, man and cybernetics – part B: cybernetics*, vol. 28, n. 4, August.
- NEVES, E. M. A. (1995). *Reconhecimento óptico de caracteres isolado baseado em suas características fundamentais*. Dissertação (Mestrado), EESC - USP, São Carlos - SP, 191 Pp., Março.
- NEUBAUER, C. (1998). Evaluation of convulational neural networks for visual recognition, *IEEE Trans. Neural Networks*, 9 (4), pp. 685-696.
- NIELSON, G. M.; HAGEN, H.; MÜLLER, H. (1997). *Scientific visualization: overviews, methodologies and techniques*, Ed. IEEE Computer Society, California.
- NIGRIN, A. (1993). *Neural Networks for pattern recognition*, A Bradford Book (The MIT Press).
- NORDIC (2001). *Wavefront analysis and adaptive optics in vision research*. <http://www.oftalmologi.gu.se/hp/research/wavefront.html> (June).
- OPTICS (2001). *Integrated wavefront sensor*. http://guernsey.et.tudelft.nl/group/project_davies.html (June).
- OSAJIMA, I. (1992). *Reconstrução de imagem digital aplicada à angiografia*. Dissertação (Mestrado), Escola de Engenharia de São Carlos - USP, São Carlos - SP, 111 Pp., Novembro.
- PAVLIDIS, T. (1982). *Algorithms for graphics and image processing*, Ed. Springer, Berlin.
- PEREIRA, A. S. (1995). *Processamento de imagens médicas utilizando a transformada de hough*, Tese (Doutorado), Instituto de Física de São Carlos - USP, São Carlos - SP, 263 Pp., Maio.
- PRAMPERO, P. S. (1998). *Combinação de classificadores para reconhecimento de padrões*, Dissertação (Mestrado), Instituto de Ciências Matemáticas e Computação - USP, São Carlos - SP, 84 Pp., março.
- PROCEEDING'96 (1996). *Proceedings computer vision and pattern recognition*, S. Francisco (CA), June 18-20, IEEE Computer Society Conference.
- RICH, E. & KNIGHT, K. (1994). *Inteligência Artificial*, Makron Books do Brasil Editora Ltda., 722 Pp.
- RIDDER, D. et al (1999). The applicability of neural networks to non-linear image processing, *Pattern Anal. Application*, 2 (2), pp. 111-128.
- ROORDA, A. (2000). *Adaptive optics for the human eye*. Center for Visualization, University of Rochester, http://www.cvs.rochester.edu/~sroorda/ao_research.htm (April).
- SCHALKOFF, R. (1992). *Pattern recognition: statistical, structural and neural approaches*. Ed. Wiley.
- SCHNEIDER, H & FRANK, P. (1996). Observer-based supervision and fault-detection in robots using nonlinear and fuzzy logic residual evaluation, *IEEE Transactions on Control Systems Technology*, vol.4, 3, pp. 274-282.
- SCOTTISH UNIVERSITIES S. S. IN PHYSICS (1991). *Pattern recognition and image processing in physics /S. A.*, Ed. Bristol, 366 Pp.
- SEGATTO, E. C. (1999). *Redes neurais artificiais aplicadas à proteção diferencial de transformadores de potência*, Dissertação (Mestrado), Escola de Engenharia de São Carlos – USP, São Carlos, 108 p.

- SERRA, J. P. F. (1988). *Image analysis and mathematical morphology*, volume 2: theoretical advances. Academic Press, London, 411 Pp.
- SHARKEY, A. J. C. (1996). On combining artificial neural nets, *Connection Science*, v. 8, n.3, pp. 299-313.
- SHIRVAIKAR, M. V. & TRIVEDI, M. M. (1995). A neural network filter to detect small targets in high clutter backgrounds, *IEEE Transactions on Neural Networks*, vol. 6, n.1.
- SILVA, E. P. et al. (1996). Desenvolvimento de um instrumento para medida da topografia da córnea, *Anais do IX SIBGRAPI*, Caxambu – MG, pp. 361-362, outubro.
- SILVA, R. E. (2001). *Um simulador para o teste de acuidade visual usando ray-tracing*. Dissertação (Mestrado), ICMC - USP, São Carlos - SP, 90 Pp.
- SRINIVASAN, V. et al. (1994). Edge detection using a neural network. *Pattern Recognition*, vol. 27, n. 12.
- THIBOS, L. N. (2000). The prospects for perfect vision. *Journal of refractive surgery*, vol. 16, pp.540-546, September/October.
- TRAINA, A. J. M. (1991). *Estudo e implementação do software dedicado para um sistema de visualização de imagens*. Tese (Doutorado), IFQSC - USP, São Carlos - SP, 124 Pp., julho.
- TRONCO, M. L. (1999). *Sistemas de reconhecimento de imagens baseado no modelo GSN de rede neural*, Tese (Doutorado), EESC – USP, São Carlos, 244 p.
- VAILLANT, R. et al. (1994). Original approach for the localization of objects in images, *IEE Proceedings – Visualization, Image Signal Processing*, vol. 141, n. 4.
- VALENTE, C. M. O. (1999). *Fixação de objetos de formato geométrico desconhecido utilizando redes neurais artificiais*, Dissertação (Mestrado), EESC-USP, São Carlos - SP, 120 Pp.

ANEXO 1 – COMPARAÇÃO DOS VETORES DE CARACTERÍSTICAS

Este anexo apresenta os gráficos comparativos dos diferentes vetores de características gerados para os pares de imagens de um mesmo olho, discutidos na Seção 6.3.3. As Figuras AN1.1, AN1.2 e AN1.3 apresentam os gráficos da diferença (em módulo) entre os atributos dos vetores de características obtidos com a PCA (6.400 atributos) para as três imagens do olho do paciente 24o68D, e as Figuras AN1.4, AN1.5 e AN1.6 apresenta os gráficos correspondentes obtidos com os vetores de características gerados pela transformada de Gabor (200 atributos) para o mesmo olho.

Figura AN1.1 – Diferença entre os vetores de características obtidos com PCA – imagens 1 e 2.

Figura AN1.2 – Diferença entre os vetores de características obtidos com PCA – imagens 2 e 3.

Figura AN1.3 – Diferença entre os vetores de características obtidos com PCA – imagens 3 e 1.

Figura AN1.4 – Diferença entre os vetores de características obtidos por Gabor – imagens 1 e 2.

Figura AN1.5 – Diferença entre os vetores de características obtidos por Gabor – imagens 2 e 3.

Figura AN1.6 – Diferença entre os vetores de características obtidos por Gabor – imagens 3 e 1.

ANEXO 2 – GRÁFICOS DOS RESULTADOS DOS EXPERIMENTOS COM GABOR E SVM

Este anexo apresenta os gráficos dos resultados obtidos nos experimentos com o algoritmo SVM descritos na Seção 7.3, que utilizam como entrada os atributos extraídos pela transformada *wavelet* de Gabor. São apresentados, separados por classe, o número de exemplos (padrões) e o erro médio total relativos aos conjuntos de teste (10 partições) para cada uma das medidas S , C e A .

As Figuras AN2.1 e AN2.2 mostram os resultados referentes à medida S , sendo que a Figura AN2.2 apresenta os resultados com propagação do erro. Os resultados referentes às medidas C e A são mostrados, respectivamente, nas Figuras AN2.3 e AN2.4 (com propagação do erro), e AN2.5 e AN2.6 (com propagação do erro).

As Figuras AN2.7 e AN2.8 mostram os resultados referentes à medida S , após o balanceamento artificial, sem e com a propagação do erro, respectivamente. Os resultados referentes à medida C após o balanceamento artificial são mostrados nas Figuras AN2.9 e AN2.10 (com propagação do erro), e os resultados com balanceamento artificial referentes à medida A são exibidos nas Figuras AN2.11 e AN2.12 (com propagação do erro).

Figura AN2.1 – Porcentagens do número de padrões e do erro médio total para a medida *S*.

Figura AN2.2 – Porcentagens do número de padrões e do erro médio total para a medida *S* com propagação do erro.

Figura AN2.3 – Porcentagem do número de padrões e do erro médio total para a medida C.

Figura AN2.4 – Porcentagens do número de padrões e do erro médio total para a medida C com propagação do erro.

Figura AN2.5 – Porcentagem do número de padrões e do erro médio total para a medida A.

Figura AN2.6 – Porcentagens do número de padrões e do erro médio total para a medida A com propagação do erro.

Figura AN2.7 – Porcentagem do número de padrões e do erro médio total para a medida *S* com balanceamento artificial

Figura AN2.8 – Porcentagens do número de padrões e do erro médio total para a medida *S* com balanceamento artificial e propagação do erro.

Figura AN2.9 – Porcentagem do número de padrões e do erro médio total para a medida *C* com balanceamento artificial

Figura AN2.10 – Porcentagens do número de padrões e do erro médio total para a medida *C* com balanceamento artificial e propagação do erro.

Figura AN2.11 – Percentagens do número de padrões e do erro médio total para a medida A com balanceamento artificial

Figura AN2.12 – Percentagens do número de padrões e do erro médio total para a medida A com balanceamento artificial e propagação do erro.