

SERVIÇO DE PÓS-GRADUAÇÃO DO ICMC-USP

Data de Depósito:

Assinatura : _____

Recuperação de imagens por conteúdo através de análise multiresolução por *Wavelets*

César Armando Beltrán Castañón

Orientação: *Profa. Dra. Agma Juci M. Traina*

Dissertação apresentada ao Instituto de Ciências Matemáticas e de Computação - ICMC-USP, para obtenção do título de Mestre em Ciências de Computação e Matemática Computacional.

USP - São Carlos
Fevereiro de 2003

Recuperação de imagens por conteúdo através de análise multiresolução por *Wavelets*

César Armando Beltrán Castañón

As minhas adoradas, Alejandra e Ledy,
por serem a razão de minha vida;
aos meus pais, Esther e Pedro,
neles encontro meu porto seguro;
aos meus irmãos, Norman e Anibal,
por serem meus melhores amigos.

Agradecimentos

A minha orientadora, Professora Dra. Agma J. Traina, por sua grande ajuda na orientação, suporte e fortalecimento. Ela me forneceu não só o conhecimento técnico, mas também a atitude rigorosa na pesquisa. Sinto-me afortunado por ter sido orientando dela.

A todo o pessoal do Grupo de Bases de Dados e Imagens (GBDI-ICMC), pela acolhida, abertura e disponibilidade em me ajudar.

A minha sogra, “mamá Trina”, minhas cunhadas e cunhados, sobrinhos e sobrinhas, que me apoiaram e proporcionaram a tranqüilidade necessária, cuidando de minha filha e minha esposa, durante minha permanência em São Carlos.

Ao meu tio, Edwin, às minhas tias, Lucy e Ruth, e aos meus primos Medalith, Luis Alberto, Junior, Nandy e Fallon; pela motivação que sempre recibi deles.

Aos amigos do alojamento, especialmente ao Clever; meus colegas da turma e da nova turma, com quem compartilhei idéias e momentos de companheirismo.

Aos colegas e amigos da Sociedade Peruana de Computação, especialmente ao Ernesto e sua esposa Govy, pelos conselhos e apoio durante o tempo dos estudos de mestrado.

À Beth, Laura e Ana, pela atenção dispensada.

Ao CNPq, pelo apoio financeiro.

Resumo

Os sistemas de recuperação de imagens por conteúdo (CBIR - *Content-based Image Retrieval*) possuem a habilidade de retornar imagens utilizando como chave de busca outras imagens. Considerando uma imagem de consulta, o foco de um sistema CBIR é pesquisar no banco de dados as n imagens mais similares à imagem de consulta de acordo com um critério dado. Este trabalho de pesquisa foi direcionado na geração de vetores de características para um sistema CBIR considerando bancos de imagens médicas, para propiciar tal tipo de consulta.

Um vetor de características é uma representação numérica sucinta de uma imagem ou parte dela, descrevendo seus detalhes mais representativos. O vetor de características é um vetor n -dimensional contendo esses valores. Essa nova representação da imagem pode ser armazenada em uma base de dados, e assim, agilizar o processo de recuperação de imagens. Uma abordagem alternativa para caracterizar imagens para um sistema CBIR é a transformação do domínio. A principal vantagem de uma transformação é sua efetiva caracterização das propriedades locais da imagem.

Recentemente, pesquisadores das áreas de matemática aplicada e de processamento de sinal desenvolveram técnicas práticas de *wavelet* para a representação multiescala e análise de sinais. Estas novas ferramentas diferenciam-se das tradicionais técnicas de Fourier pela forma de localizar a informação no plano tempo-freqüência; basicamente, elas têm a capacidade de mudar de uma resolução para outra, o que faz delas especialmente adequadas para a análise de sinais não estacionários. A transformada *wavelet* consiste de um conjunto de funções base que representa o sinal em diferentes bandas de freqüência, cada uma com resoluções distintas correspondentes a cada escala. Estas foram aplicadas com sucesso na compressão, melhoria, análise, classificação, caracterização e recuperação de imagens. Uma das áreas beneficiadas, onde essas propriedades têm encontrado grande relevância, é a área médica, através da representação e descrição de imagens médicas.

Este trabalho descreve uma abordagem para um banco de imagens médicas, que é orientada à extração de características para um sistema CBIR baseada na decomposição multiresolução de *wavelets* utilizando os filtros de Daubechies e Gabor. Essas novas características de imagens foram também testadas utilizando uma estrutura de indexação métrica *Slim-tree*. Assim, pode-se aumentar o alcance semântico do sistema cbPACS (*Content-Based Picture Archiving and Communication Systems*), atualmente em desenvolvimento conjunto entre o Grupo de Bases de Dados e Imagens do ICMC-USP e o Centro de Ciências de Imagens e Física Médica do Hospital das Clínicas de Ribeirão Preto-USP.

Abstract

Content-based image retrieval (CBIR) refers to the ability to retrieve images on the basis of the image content. Given a query image, the goal of a CBIR system is to search the database and return the n most similar (close) ones to the query image according to a given criteria. Our research addresses the generation of feature vectors of a CBIR system for medical image databases.

A feature vector is a numeric representation of an image or part of it over its representative aspects. The feature vector is a n -dimensional vector organizing such values. This new image representation can be stored into a database and allow a fast image retrieval. An alternative for image characterization for a CBIR system is the domain transform. The principal advantage of a transform is its effective characterization for their local image properties.

In the past few years, researches in applied mathematics and signal processing have developed practical *wavelet* methods for the multiscale representation and analysis of signals. These new tools differ from the traditional Fourier techniques by the way in which they localize the information in the time-frequency plane; in particular, they are capable of trading one type of resolution for the other, which makes them especially suitable for the analysis of non-stationary signals. The *wavelet* transform is a set of basis functions that represents signals in different frequency bands, each one with a resolution matching its scale. They have been successfully applied to image compression, enhancements, analysis, classifications, characterization and retrieval. One privileged area of application where these properties have been found to be relevant is medical imaging.

In this work we describe an approach to CBIR for medical image databases focused on feature extraction based on multiresolution *wavelets* decomposition, taking advantage of the Daubechies and Gabor. Fundamental to our approach is how images are characterized, such that the retrieval procedure can bring similar images within the domain of interest, using a metric structure indexing, like the *Slim-tree*. Thus, it increased the semantic capability of the cbPACS (*Content-Based Picture Archiving and Comunication Systems*), currently in joined developing between the Database and Image Group of the ICMC-USP and the Science Center for Images and Physical Medic of the Clinics Hospital of Ribeirão Preto-USP.

Índice

1	Introdução	1
1.1	Considerações Iniciais	1
1.2	Motivação	3
1.3	Objetivos	5
1.4	Contribuições	5
1.5	Organização do Trabalho	6
2	Caracterização de Imagens	7
2.1	Considerações Iniciais	7
2.2	Definindo uma Imagem	8
2.3	Atributos das Imagens	9
2.3.1	Cor	9
2.3.2	Textura	11
2.3.3	Forma	12
2.3.4	Resumo de Características de Baixo Nível	13
2.4	Exemplos de Características Extraídas de Imagens 2D	14
2.4.1	Estatística Descritiva para Imagens	15
2.4.2	Descrição de Regiões	16
2.5	Considerações Finais	23
3	Introdução à Teoria de Wavelets	25
3.1	Considerações Iniciais	25
3.2	Introdução Histórica	26
3.3	Transformada de Fourier	26
3.4	Transformada Wavelet	27
3.4.1	Métodos Multiresolução	28
3.4.2	Transformada Wavelet de Haar	32
3.4.3	Transformada Wavelet de Daubechies	32
3.4.4	Extensões para Dimensões Maiores	35
3.5	Transformada Wavelet de Gabor	37
3.5.1	Funções Gabor e Wavelets	38
3.6	Considerações Finais	40

4 Wavelets e Extração de Características	43
4.1 Considerações Iniciais	43
4.2 Similaridade Semântica das Imagens	43
4.3 Propriedades e Características das Imagens Médicas	44
4.4 Escolha da Wavelet Analisadora	45
4.5 Geração do Espaço de Características	47
4.6 Medidas de Distância	48
4.7 Análise e Avaliação Visual	49
4.8 Considerações Finais	50
5 Experimentos e Resultados	51
5.1 Introdução	51
5.2 Visão Geral	51
5.3 Conjunto de Dados	52
5.4 Grupos de Teste	54
5.4.1 Grupo Básico	54
5.4.2 Grupo Intermediário	54
5.4.3 Grupo Maior	56
5.5 Características Baseadas na Distribuição da Cor	57
5.5.1 Motivação	57
5.5.2 Algoritmo de Decomposição	57
5.5.3 Geração dos Vectors de Características	58
5.5.4 Visualização dos Vectors de Características	59
5.5.5 Aplicando outros Extratores	60
5.6 Características Baseadas na Textura da Imagem	62
5.6.1 Motivação	62
5.6.2 Extração do Vetor de Características	62
5.6.3 Desempenho do Extractor	63
5.6.4 Mudança na Medida de Distância	63
5.7 Avaliação de Resultados no Sistema Protótipo	66
5.7.1 Tempo de Geração de Características	66
5.7.2 Tempo de Resposta	67
5.8 Utilizando o Método de Acesso Métrico <i>Slim-tree</i>	68
5.9 Considerações Finais	70
6 Conclusões	71
6.1 Considerações Finais	71
6.2 Linhas de Futuras Pesquisas	72
Referências Bibliográficas	73
A Revisão de Álgebra Linear	83
A.1 Espaço Vetorial ou Linear	83
A.2 Bases e Dimensões	83
A.3 Corpo	84
A.4 Produto Interno e Ortogonalidade	84
A.5 Normas e Normalização	85
B Medidas de Distância	87

C Descrição do Protótipo de Sistema de Consulta	89
C.1 Transformação de Imagens	89
C.2 Geração do Espaço de Características	90
C.3 Recuperação de Imagens	90
C.4 Sistema de Indexação de dados	90
C.5 Visualização de Consultas	91

Lista de Figuras

1.1	Processo de geração de assinatura aplicando transformada <i>wavelet</i>	3
2.1	Imagen original	10
2.2	Imagen saturada	10
2.3	Histograma da imagen original	10
2.4	Histograma da imagen saturada	10
2.5	Imagenes com textura (Wang, 2000)	11
2.6	Formas de imagens (da Fontoura Costa and Cesar, 2000)	12
2.7	As quatro orientações usadas para construir as matrizes de co-ocorrência (Theodoridis, 1999)	17
3.1	Uma função contínua $f(x)$ (linhas pontilhadas) e a função de aproximação caixa (linhas sólidas) em duas resoluções. Na esquerda a aproximação grosseira $A_0f(x)$ e na direita é a aproximação de maior resolução $A_1f(x)$ (Davis and Nosratinia, 1998).	29
3.2	Geração do espaço superior em função de $\phi(x)$ e $\psi(x)$	30
3.3	Algumas análise <i>wavelets</i> . Primeira fila: <i>wavelets</i> pai, $\phi(x)$. Segunda fila: <i>wavelets</i> mãe, $\psi(x)$ (Daubechies, 1992)	33
3.4	Processo de decomposição ao aplicar a transformada <i>wavelet</i> em um domínio 2D (Albuz et al., 2001)	36
3.5	Transformada <i>wavelet</i> sobre um domínio 2D. Inicia-se com a imagen original e, após transformala, obter uma imagen com dois níveis de decomposição.	36
3.6	Componentes real e imaginário de $\psi(x, y)$	38
3.7	(a) Conjunto de <i>wavelets</i> Gabor (comprimento de banda octal de 1.5) (b) sua cobertura no plano de freqüência espacial (Lee, 1996)	39
3.8	Filtros de resposta ortogonais, após aplicar a técnica de eliminação de redundância (Manjunath and Ma, 1996)	40
4.1	Imagen com áreas de textura que identificam regiões distintas	45
5.1	Processo geral de extração de características	52
5.2	Imagenes de RM e CT (percebe-se a diferença de tons de cinza)	53
5.3	Seqüência de imagens RM com espessura distinta nos cortes	53
5.4	Exemplos de imagens da base utilizada	54
5.5	Imagenes do <i>GrupoBásico</i> de teste, cada imagen é uma representativa de cada grupo	55

5.6	Imagens do <i>GrupoIntermediário</i> de teste, cada imagem é uma representativa de cada grupo	55
5.7	(a) Imagem original, (b) Decomposição da imagem após três níveis de transformação com filtros Daubechies-4, o contraste e brilho da imagem foram mudados com a finalidade de ressaltar os detalhes.	58
5.8	Geração do vetor de características de uma imagem baseada nos subespaços de <i>wavelets</i> considerando três níveis de decomposição e a função $E(F_n)$ corresponde ao extrator aplicado sobre o subespaço n	59
5.9	Visualização no FastMapDB correspondente a cinco grupos de imagens utilizando como extrator a energia e filtros de <i>wavelets</i> Daubechies-4	60
5.10	Grafo de precisão e revocação das <i>wavelets</i> Daubechies-4	60
5.11	Visualização correspondente aos extractores (a) média e (b) entropia, com filtros de <i>wavelets</i> Daubechies-4	61
5.12	Visualização correspondente ao <i>GrupoBásico</i> aplicando filtros de <i>wavelets</i> Haar e extractores (a) energia, (b) média e (c) entropia.	61
5.13	Visualização do extrator Gabor usando como métrica a distância Euclideana para (a) grupo de teste <i>TesteBásico</i> e (b) grupo de teste <i>TesteIntermediário</i>	63
5.14	Grafo de precisão e revocação para o extrator de Gabor com a distância Euclideana como métrica, sobre o grupo de teste <i>GrupoIntermediário</i>	64
5.15	Grafo de precisão e revocação dos quatro extractores (<i>média, energia, entropia e gabor</i>), distância Euclideana e grupo de teste <i>GrupoIntermediário</i>	64
5.16	Visualização do extrator Gabor usando como métrica a distância definida por Manjunath (Manjunath and Ma, 1996) comparando com os outros extractores, grupo de teste <i>TesteIntermediário</i>	65
5.17	Médias dos tempos de resposta para diferentes consultas dos k -vizinhos mais próximos sobre a base de 5631 vetores de características utilizando a <i>Slim-tree</i> e busca seqüencial.	69
5.18	Grafo que mostra o número de comparações que são realizados na <i>Slim-tree</i> para recuperar k -vizinhos mais próximos, junto à estrutura seqüencial (resposta constante) para um espaço composto de 5631 vetores de características.	69
C.1	Esquema do sistema de consulta por similaridade	90
C.2	Tela principal do sistema	91
C.3	Consulta com Gabor-2-6 utilizando distância Euclideana	92
C.4	Consulta com Gabor-2-6 utilizando distância Euclideana Normalizada	92
C.5	Consulta com Gabor-4-6 utilizando distância Euclideana	93
C.6	Consulta com Gabor-4-6 utilizando distância Euclideana Normalizada	93
C.7	Consulta com Gabor-2-6 utilizando distância Euclideana Normalizada, procurando uma imagem que não pertence ao banco de imagens	94
C.8	Consulta com Gabor-4-6 utilizando distância Euclideana Normalizada, procurando uma imagem que não pertence ao banco de imagens	94
C.9	Consulta com Gabor-4-6 utilizando distância Euclideana Normalizada de uma imagem aleatória	95
C.10	Consulta com Gabor-4-6 utilizando distância Euclideana Normalizada de uma imagem aleatória	95

Lista de Tabelas

2.1	Descrição de tipos de características de baixo nível (Vailaya, 2000)	14
2.2	Comparação de sistemas existentes (Albuz et al., 2001)	14
5.1	Conjunto de imagens de teste - <i>GrupoBásico</i>	55
5.2	Conjunto de imagens de teste - <i>GrupoIntermediário</i>	56
5.3	Conjunto de imagens de teste - <i>GrupoMaior</i>	56
5.4	Tempo de geração de características baseadas em filtros de Gabor com distintos parâmetros (exemplo 2 – 6 = 2 escalas - 6 orientações).	66
5.5	Tempo médio de resposta de consulta no sistema protótipo usando indexação seqüencial junto à função Euclideana, para 5631 imagens.	67
5.6	Resultados do tempo médio de resposta das consultas feitas no protótipo com uma outra distância (Euclideana normalizada), para 5631 imagens	68

Introdução

1.1 Considerações Iniciais

Nos últimos anos tem-se visto um rápido aumento no tamanho das coleções de imagens digitais. A cada dia centros militares e sociais vêm gerando *gigabytes* de imagens. Porém, ao mesmo tempo, imensas quantidades de informação estão ficando sem utilidade, pois não é possível ter acesso a essa informação se, previamente, ela não houver sido organizada para permitir busca e recuperação eficientes de dados. No começo dos anos 90, devido ao surgimento de coleções de imagens em grande escala, surgiram dois grandes problemas para o tratamento dessas imagens. O primeiro é o esforço para se fazer as anotações manuais. O segundo problema, e o mais importante, refere-se ao entendimento do valioso conteúdo das imagens e a subjetividade da percepção humana. Para superar essas dificuldades foi proposta a recuperação de imagens por conteúdo (“*Content-Based Image Retrieval - CBIR*”), que, em vez de utilizar as anotações manuais mediante palavras chaves, as imagens podem ser indexadas mediante seu conteúdo visual como, por exemplo, cor, textura, forma, etc.(Yong Rui, 1999).

São três as bases fundamentais da recuperação de imagens por conteúdo: extração de características visuais, indexação multidimensional e projeto de sistemas de recuperação (Yong Rui, 1999). A extração de características (que reflete a semântica inerente da imagem) é a base da recuperação de imagens por conteúdo. Em um sentido mais amplo, as características podem estar baseadas no conteúdo textual (chaves, anotações, etc.) ou visual (cor, textura, forma, relacionamento entre objetos, etc.). Dentro do escopo das características visuais, estas podem ser classificadas como características gerais e características específicas de domínio. As primeiras,

incluem características de cor, textura e forma; enquanto as últimas são dependentes da aplicação, incluindo características específicas, por exemplo, faces humanas e impressões digitais em sistemas de reconhecimento de criminosos. As características de domínio específico são melhor cobertas em artigos de reconhecimento de padrões e podem incluir temas referentes ao domínio do conhecimento. Neste trabalho, a extração de características refere-se às características gerais, que podem ser usadas em muitas aplicações. O enfoque deste trabalho é para o tratamento de imagens médicas.

Com o desenvolvimento de computadores com maiores capacidades de processamento e armazenamento, a aquisição digital de imagens médicas tornou-se bastante popular nos últimos anos. Além dos sistemas tradicionais de raios-X e mamografia, técnicas mais recentes tais como imagens de ressonância magnética (*Magnetic Resonance Imaging -MRI*) e tomografia computadorizada (*Computer Tomography - CT*), podem produzir muitas imagens por exame. Logo, a cada ano, um hospital pode produzir muitos *terabytes* de imagens médicas digitais (Weiss, 2002). Atualmente o armazenamento já não é um problema grande por causa da diminuição do custo dos meios de armazenamento secundário. Mas, o uso efetivo do banco de dados de imagens médicas de grande escala ainda é um desafio para os pesquisadores nesta área (Wang, 2001).

Os sistemas de informação hospitalares (SIH) são projetados para organizar as informações textuais de pacientes (nome, idade, tratamento, etc.) e em alguns casos permitem a associação das informações textuais com informações gráficas, provenientes de exames como Raio-X, tomografia (CT ou RM), ultra-som, etc. Recentemente, com a introdução dos sistemas PACS (*Picture Archiving and Communication Systems*), cresceu o interesse por integrar em um só sistema todas as informações dos pacientes (textos, imagens, gráficos temporais, etc). Em sistemas PACS, pode-se recuperar também informações baseadas apenas no conteúdo das próprias informações. Por exemplo, pode-se querer recuperar “todas as imagens obtidas em exames de CT de coração similares a uma imagem dada”. Nesse caso, o cruzamento das informações é independente das chaves de busca tradicionais baseadas em informações numéricas e textuais, que associam exames e pacientes em um SIH tradicional. A facilidade de recuperação de dados baseada em seu conteúdo é um dos pontos altos dos sistemas PACS (Korn et al., 1996) (de Lima et al., 1998). Através dela pode-se encontrar registros de pacientes cujas imagens associadas apresentam aspectos de similaridade (utilizando padrões gráficos) em comum, além de obter-se correlações utilizando sintomas previamente registrados, tratamentos efetuados, etc. (Marsh, 1997). Um sistema PACS pode ser uma ferramenta bastante eficaz no auxílio ao diagnóstico médico.

Um sistema PACS deve então ser capaz de responder consultas por similaridade, isto é, questões do tipo: “quais são as 10 imagens mais próximas de uma dada imagem?” ou “quais são as imagens que diferem de até 5 unidades de uma imagem padrão?”. Nota-se que tanto as imagens quanto a função de dissimilaridade (distância) entre elas devem ser definidas.

Uma alternativa para indexar imagens por conteúdo é a indexação pelo domínio de transformada. A principal vantagem esperada de uma transformação é a efetiva caracterização das propriedades locais da imagem. Além disso, as transformações podem ser utilizadas para a compressão das imagens e a indexação baseada em dados comprimidos incrementa a eficiência do armazenamento e a performance do sistema.

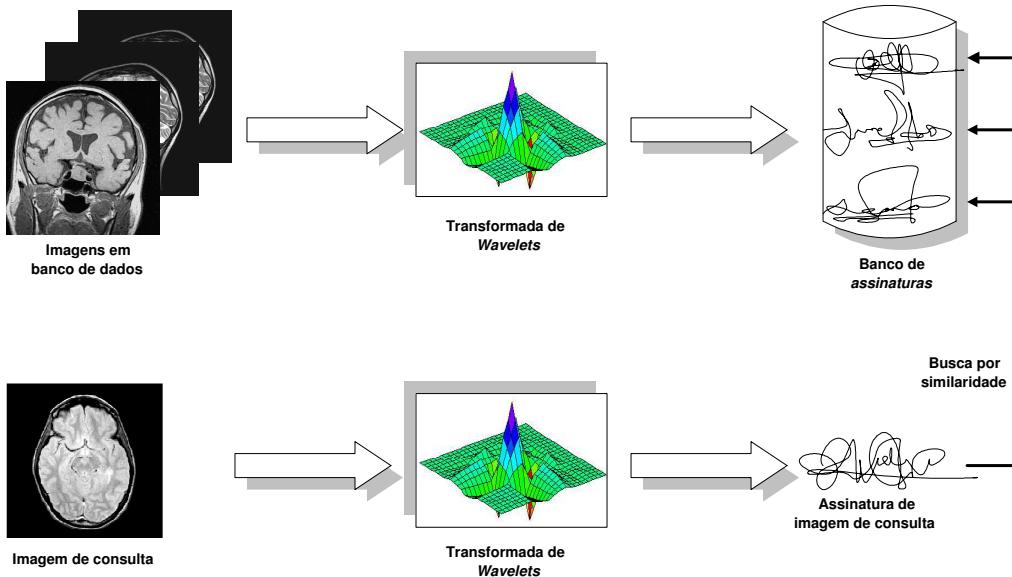


Figura 1.1: Processo de geração de assinatura aplicando transformada *wavelet*

A transformada *wavelet* é uma técnica de processamento de sinais que decompõe um sinal em diferentes sub-bandas de freqüência. Este modelo pode ser generalizado para sinais d -dimensionais nos quais uma transformada 1-dimensional pode ser aplicada múltiplas vezes. Para cada objeto, as características extraídas formam um vetor de características que pode ser representado mediante um ponto no espaço de características d -dimensional, na figura 1.1 o vetor de características é representado como a assinatura da imagem (Jacobs et al., 1995). Uma base de dados espacial será a coleção de tais pontos. A transformada *wavelets* pode ser aplicada sobre os objetos para gerar os vetores de características (espaço de características) (Sheikholeslami et al., 1998).

1.2 Motivação

Ao gerenciar e manipular informações complexas e volumosas, imagens de exames médicos por exemplo, um dos problemas a serem tratados é o de como armazenar e recuperar tais informações de maneira rápida e precisa. Uma técnica usual é a extração de características (parâmetros) da imagem que idealmente possibilitem a identificação da mesma de forma inequívoca, ou com a menor ambigüidade possível. Muitas vezes esse processo é baseado na redução da dimen-

sionalidade dos dados: parâmetros são extraídos das imagens e a seguir são utilizados para indexá-las agilizando sua busca e recuperação.

As estruturas de indexação são ferramentas fundamentais nos sistemas de gerenciamento de banco de dados. As primeiras estruturas de indexação desenvolvidas, como as estruturas ISAM (“*Indexed Sequential Access Method*”) são voltadas à indexação de números e pequenas cadeias de caracteres, sendo que os gerenciadores de dados comerciais atuais as utilizam maciçamente. Essas estruturas não são úteis para dados complexos, como imagens, sons, vídeo, séries temporais ou estruturas moleculares, que são tipicamente conjuntos de dados de altas dimensões ou mesmo adimensionais. Para o tratamento de dados espaciais com poucas dimensões, já existem diversas estruturas de indexação adequadas, genericamente associadas a Métodos de Acesso Espaciais - SAM (“*Spatial Access Methods*”), cuja pesquisa iniciou-se com o trabalho pioneiro sobre as *R-Trees* (Guttman, 1984) e prosseguiu com numerosas variações e adaptações (Sellis et al., 1987) (Beckmann et al., 1990) (Hellerstein et al., 1995) (Papadias et al., 1995) (ou veja (Gaede and Günther, 1998) para um apanhado geral sobre estruturas espaciais). No entanto, para dados em espaços métricos adimensionais, nenhuma dessas estruturas é aplicável.

Estruturas de dados para espaços métricos (que englobam tanto dados espaciais com dimensão definida quanto dados adimensionais) que abordam essa necessidade estão começando a ser estudadas, embora existam ainda poucos trabalhos na área. Exemplos de dados métricos são imagens, impressões digitais, palavras de um dicionário, com funções específicas de dissimilaridade associadas a elas. As principais estruturas métricas existentes são: a “*fixed query tree*” (Baeza-Yates et al., 1994), a *mvp-tree* (*multi-vantage point tree*) (Bozkaya and Özsoyoglu, 1997) (Bozkaya and Özsoyoglu, 1999), a *vp-tree* (*vantage-point* (Chiueh, 1994)), a GNAT (Brin, 1995), a *M-tree* (Ciaccia et al., 1997) e a *Slim-tree* (Traina et al., 2000). Dessas estruturas, apenas a *M-tree* e a *Slim-tree* são dinâmicas, ou seja, permitem a inclusão e remoção de novos objetos depois da árvore criada – as demais estruturas são criadas em uma única operação sobre o conjunto de dados, que torna-se imutável depois da criação da árvore. Além disso, ambas *M-tree* e *Slim-tree* são escaláveis e mantêm eficiência quando armazenadas em memória secundária.

Neste projeto, as características das imagens serão extraídas através de *wavelets* gerando vetores de características. A seguir as imagens serão indexadas e recuperadas utilizando a estrutura *Slim-tree*. Dessa forma, será incluída uma nova funcionalidade ao sistema cbPACS (<http://www.cci.fmrp.usp.br>) que está atualmente em desenvolvimento conjunto entre o Grupo de Bases de Dados e Imagens GBDI do ICMC-USP e o Centro de Ciências de Imagens e Física Médica do Hospital das Clínicas da Faculdade de Medicina de Ribeirão Preto-USP.

1.3 Objetivos

O objetivo deste trabalho foi desenvolver um conjunto de técnicas para a extração de características de imagens médicas, para recuperação de imagens por conteúdo, através da análise multiresolução por *wavelets*.

Esse conjunto de técnicas implicou em desenvolver um **extractor de características** (utilizando análise multiresolução por transformada *wavelet*) e a procura de uma métrica que possibilite a comparação mais precisa entre duas imagens. Através dessas características, as imagens podem ser organizadas no sistema cbPACS, possibilitando sua recuperação por conteúdo. Dessa forma, o sistema cbPACS pode aumentar o alcance semântico e conseguir uma nova funcionalidade para atender as consultas por similaridade, buscando satisfazer às necessidades de buscas semânticas sobre as imagens.

1.4 Contribuições

Além da revisão bibliográfica de alguns métodos de caracterização de imagens, na busca por novos métodos eficientes de recuperação de imagens por conteúdo, foram estudados vários algoritmos, implementados e testados alguns métodos. Como resultado, serão apresentadas as seguintes contribuições:

- Descrevem-se maneiras de se aplicar as transformadas *wavelets* no processo de caracterização de imagens (Capítulos 4 e 5).
- Analisa-se os resultados obtidos, definindo as distintas aplicabilidades das funções base de *wavelet*, em distintos domínios de imagens médicas.
- Propõe-se o tratamento de textura na análise das imagens médicas mediante filtros de Gabor.
- Propõe-se a utilização de uma função distância específica para o cálculo da similaridade das imagens através de texturas.
- Avaliação visual das características extraídas, aproveitando a ferramenta FastMapDB desenvolvida no GBDI.
- Desenvolvimento de um sistema protótipo para extração de características de imagens por *wavelets* (MultiWaveMed).

1.5 Organização do Trabalho

Este primeiro capítulo apresentou as considerações iniciais e o contexto no qual se insere este trabalho, a motivação para o seu desenvolvimento e os objetivos a serem alcançados.

O Capítulo 2 mostra uma visão da extração de características expondo as principais técnicas que são usadas atualmente, com especial ênfase nas características de cor e textura, nas quais está baseado o presente trabalho. Este capítulo está relacionado com técnicas utilizadas em reconhecimento de padrões da área de visão computacional.

O Capítulo 3 introduz conceitos relacionados à teoria das *wavelets* e suas aplicações na caracterização de imagens. A abordagem da transformada *wavelet* inicia com o enfoque da análise de multiresolução, culminando com a introdução à teoria dos filtros de Gabor, os quais são utilizados na caracterização textural das imagens.

O Capítulo 4 apresenta algumas considerações no desenvolvimento de sistemas CBIR, assim como descreve o estado da arte referente a extração de características e indexação de imagens médicas aplicando *wavelets*.

No Capítulo 5 são descritos e analisados os resultados obtidos visando avaliar e criar métodos para proceder com caracterização de imagens médicas.

Finalmente as conclusões e a descrição de possíveis trabalhos futuros que poderão ser implementados como continuação dessa pesquisa são apresentado no Capítulo 6.

O Apêndice A contém uma breve introdução à álgebra linear, posto que o processo de multiresolução da transformada *wavelet* está baseado nesta teoria. Algumas medidas de distância são apresentados no Apêndice B. Finalmente o Apêndice C apresenta uma breve descrição do protótipo de sistema desenvolvido para fazer consultas por similaridade.

Caracterização de Imagens

2.1 Considerações Iniciais

Nos últimos anos tem-se visto um rápido aumento no volume de coleções de imagens digitais. Por exemplo, um hospital de médio porte efetua geralmente procedimentos que geram imagens médicas que demandam centenas ou até milhares de *gigabytes* de espaço para armazenar tais dados anualmente. Administrar tais volumes de dados é um processo difícil e árduo, principalmente ao se considerar o aspecto subjetivo e/ou ambíguo das imagens, o que tem motivado o aumento de pesquisas relacionadas nessa área, por exemplo na recuperação de imagens por conteúdo (*Content-Based Image Retrieval - CBIR*) (Smeulders et al., 2000). Nos sistemas CBIR as imagens podem ser indexadas através do conteúdo visual como cor, textura e forma, onde as imagens são representadas mediante uma seqüência numérica denominada de *vetor de características*.

Um *vetor de características* é uma representação numérica sucinta de uma imagem ou parte da imagem (um objeto) caracterizando medidas dos aspectos representativos do objeto. O número de características dependerá do domínio de imagens e das propriedades que deseja-se caracterizar, esse número determinará a dimensão do vetor de características (vetor n -dimensional). Esta nova representação da imagem pode ser armazenada em um banco de dados que permita uma recuperação rápida da imagem, porém esta nova representação deve atender às três seguintes considerações (Loew, 2000):

- reduzir a dimensionalidade dos dados,
- ressaltar aspectos da imagem para facilitar a percepção humana,
- ser invariante às transformações da imagem.

Uma abordagem para esse problema é utilizar transformações de imagem, sendo que a transformada *wavelet* tem obtido alta aceitação pelos pesquisadores da área. O processo de transformação gera novos espaços de informação sobre os quais são aplicados medidas estatísticas com valores numéricos, os quais constituirão o vetor de características.

Além da transformada *wavelet* existem outras transformadas lineares que permitem desenvolver a mesma tarefa, mas neste trabalho utilizou-se a transformada *wavelet* (Haar, Daubechies e Gabor) sendo nosso maior interesse a caracterização textural das imagens. Entre essas outras transformadas têm-se: Transformada *Karhunen-Loève*, decomposição de valores singulares, transformada discreta de Fourier, transformada discreta de senos e cossenos e transformada Hadamar (Theodoridis, 1999).

Neste capítulo serão descritos alguns métodos estatísticos que suportam a criação do conjunto de técnicas que permitam caracterizar imagens após a aplicação da transformada *wavelets*. O conceito básico é transformar um conjunto de medições em um novo conjunto de características, o vetor de características (*feature vector*).

2.2 Definindo uma Imagem

Uma imagem digital é geralmente o resultado de um processo de discretização (*sampling*) de uma imagem de função contínua $I(x, y)$ e é armazenada no computador como um arranjo de duas dimensões $I(m, n)$ onde $m = 0, 1, \dots, N_x - 1$ e $n = 0, 1, \dots, N_y - 1$ o qual significa que é armazenado como um arranjo de dimensão $N_x \times N_y$.¹ Cada elemento (m, n) do arranjo corresponde a um pixel (elemento de imagem) da imagem, cujo brilho ou intensidade é igual a $I(m, n)$. Ainda mais, quando a intensidade $I(m, n)$ é quantizada em N_g níveis discretos (níveis de cinza), N_g é conhecido como a profundidade da imagem. Então a seqüência de níveis de cinza $I(m, n)$ pode tomar um dos valores $0, 1, \dots, N_g - 1$. A profundidade N_g é geralmente uma potência de 2 e pode ter valores grandes (por exemplo 64 ou 256). De qualquer forma, para o olho humano é difícil enxergar os detalhes e diferenças nas intensidades e na prática $N_g = 32$ ou 16 é suficiente para representar as imagens (Theodoridis, 1999).

A necessidade de gerar vetores de características dá-se pela dificuldade de poder trabalhar com a informação total da imagem. Por exemplo, para uma imagem de dimensão 64×64 , o número de pixels é de 4096. Para o caso de sistemas de recuperação de imagens por conteúdo,

¹ Esta definição é válida só para imagens de tons de cinza.

este número é ainda maior. Então a geração do vetor de características é um processo que calcula novas variáveis, a partir da imagem original $I(m, n)$. Este processo procura gerar características que mostrem uma grande quantidade de informações contidas em uma imagem (ou objetos da imagem). Conseqüentemente não será utilizado diretamente o total de informação da imagem $I(m, n)$, as características extraídas permitirão codificar adequadamente a informação relevante da imagem.

2.3 Atributos das Imagens

Muitos dos sistemas de recuperação de imagens utilizam cor, textura e forma, para representar uma imagem e a recuperação está baseada na similaridade das características derivadas delas. Embora a cor seja um atributo confiável na recuperação de imagens, situações onde a informação de cor não as discrimina bem, requerem do uso de atributos de textura e/ou forma para recuperação de imagens. Além disso, os sistemas baseados em um único atributo da imagem podem não alcançar níveis de recuperação adequados, motivo pelo qual os sistemas procuram utilizar múltiplos atributos das imagens para a sua indexação e recuperação (Vailaya, 2000).

2.3.1 Cor

As cores presentes em uma imagem possuem um papel significativo na indexação e recuperação da mesma. Existem diferentes representações de cores que incluem desde o tradicional RGB (*red, green, blue*), o mais simples modelo que mapeia diretamente as características físicas do dispositivo de exibição, até o HSI (*hue, saturation, intensity*) que reflete mais precisamente o modelo de cores para a percepção humana.

Muitos trabalhos recentes de extração de características baseadas em distribuição de cores estão concentrados nos histogramas de cor. Alguns dos trabalhos prévios incluíram a indexação da cor usando intersecção de histogramas (Swain and Ballard, 1991) e o trabalho recente de (Bueno, 2002) que monta uma estrutura baseada em histogramas de dimensão variável, nomeada histogramas métricos, sendo utilizados na recuperação de imagens similares. Os histogramas de cor são invariantes à translação e rotação das imagens, sendo que, com a normalização dos histogramas, obtém-se também a invariância à escala. De qualquer forma, os histogramas de cor não indicam a localização espacial dos pixels na imagem. Outros esquemas, como histogramas espaço-cromáticos (Cinque et al., 1999), utilizam a proporção de pixels que têm uma cor determinada na imagem junto com o baricentro e o desvio-padrão da região de cor homogênea, mas seu desempenho é pobre quando a disposição de uma cor que esteja espalhada (formando ilhas). Outros trabalhos enfocando a coerência da cor foram desenvolvidos em (Pass and Zabih, 1999) e (A.Vailaya et al., 1998).

Muitas operações podem ser feitas com os histogramas de cor, mudando parâmetros e limitares, por exemplo a figura 2.3 representa o histograma da imagem original (figura 2.1). Ao se aplicar uma operação de alongamento (*stretch*) sobre a imagem, o histograma ficaria como na figura 2.4, tendo como resultado uma imagem melhor definida (figura 2.2).

Alguns trabalhos, que utilizam a distribuição de cores na imagem para montar o vetor de características baseado nos coeficientes *wavelets* são apresentados por (Albuze et al., 2001), conseguindo diminuir o tempo de busca das imagens similares, mas a eficiência da busca não é tão boa (máximo de 35% de eficiência).

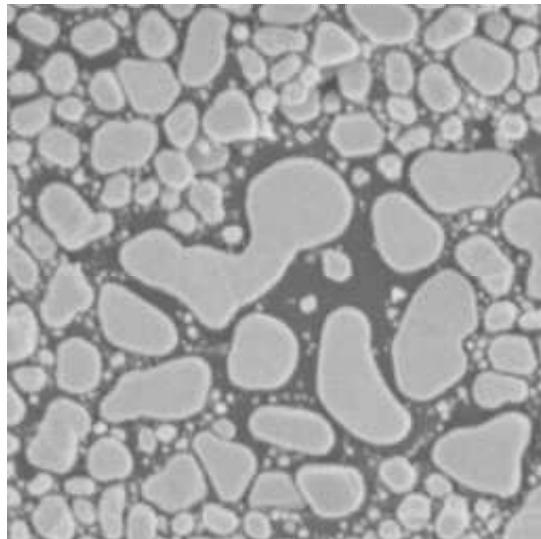


Figura 2.1: Imagem original

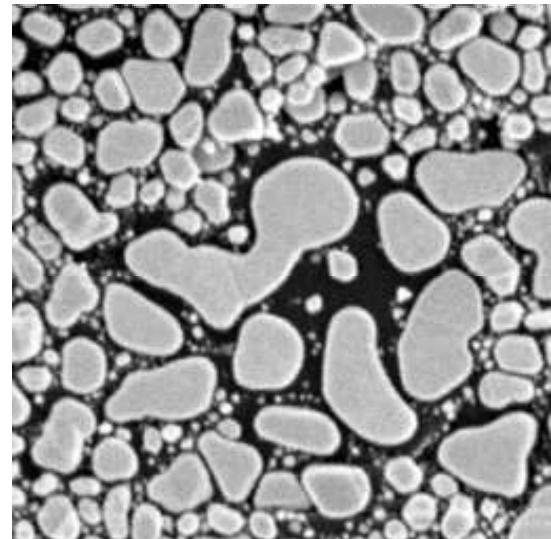


Figura 2.2: Imagem saturada

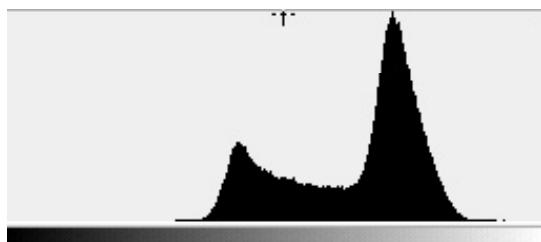


Figura 2.3: Histograma da imagem original

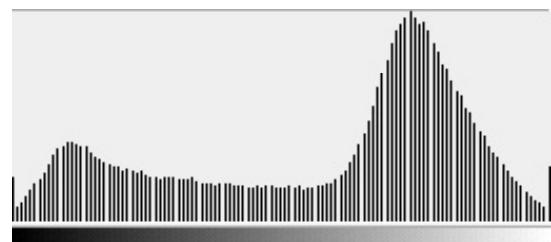


Figura 2.4: Histograma da imagem saturada

Existem sistemas comerciais muito populares que utilizam os histogramas de cores em sistemas de recuperação de imagens por conteúdo como o QBIC (Flickner et al., 1995) (<http://www.qbic.almaden.ibm.com>), Virage (<http://www.virage.com>), Excalibur (<http://www.excalib.com>), VisualSEEK (<http://disney.ctr.columbia.edu/VisualSEEk/>), NETRA (<http://maya.ece.ucsb.edu/Netra/>) entre outros.

Esta abordagem mostra algumas dificuldades pelo próprio caráter não invariante do histograma de cores. Além disso, se a imagem contém um número alto de cores, então o vetor de características será maior e a indexação de vetores com essa dimensão torna-se problemática.

2.3.2 Textura

Não existe uma definição clara de textura, mas os autores concordam em defini-la como as mudanças na intensidade da imagem que formam determinados padrões repetitivos (Tuceryan and Jain, 1993). Esses padrões podem ser o resultado de propriedades físicas da superfície do objeto (rugosidade), ou ser o resultado de diferenças de reflexão tal como a cor na superfície. A figura 2.5 mostra alguns exemplos de imagens com textura. Embora seja fácil, para as pessoas, o reconhecimento de textura, isto não acontece com procedimentos automáticos onde esta tarefa, as vezes, precisa de técnicas computacionais complexas. Nos sistemas CBIR as características de textura são muito úteis, pelo fato que a grande maioria de superfícies naturais exibem texturas.



Figura 2.5: Imagens com textura (Wang, 2000)

A análise de textura tem por objetivo estabelecer o relacionamento de vizinhança dos elementos de textura e seu posicionamento em relação aos demais (conectividade), o número de elementos por unidade espacial (densidade) e a sua regularidade (homogeneidade) (Traina, 2001).

Os modelos de textura desenvolvidos na literatura podem ser divididos nas seguintes classes (Tuceryan and Jain, 1993):

Métodos estatísticos, que definem a textura em termos da distribuição espacial dos valores de tons de cinza;

métodos geométricos, caracterizados como sendo compostos de “elementos de textura” ou primitivas;

Métodos de processamento de sinais, que usam a análise de freqüência da imagem para classificar a textura. Este último esquema inclui o uso de domínio espacial (Malik and Perona, 1990) e domínio de filtros de Fourier (Coggins and Jain, 1985), assim como o uso de filtros de Gabor e modelos de *wavelets* (Jain and Farrokhnia, 1991) (Ma and Manjunath, 1995). Estudos têm demonstrado que sistemas de textura de utilizando Gabor e multiresolução simultânea apresentam-se como uma boa técnica em sistemas de recuperação e indexação por conteúdo (Ma and Manjunath, 1995) (Picard and T.P.Minka, 1995).

No trabalho recente de (Sabino et al., 2002) aproveita-se uma nova técnica, dimensão fractal multiescala (Costa et al., 2000) (da Fontoura Costa and Cesar, 2000), para caracterizar a croma-

tina de células sanguíneas, alcançando resultados incentivadores. Nessa técnica considera-se os picos da dimensão fractal multiescala após aplicar dilatação de Minkowski-Bouligand aos tons de cinza dos núcleos, conseguindo discriminar através das diferenças nas dimensões, as células malignas das benignas.

A utilização de textura de imagens visa permitir sua segmentação em determinadas regiões, que possuam a mesma textura. Depois de definir as regiões, pode-se utilizar retângulos que as envolvam (*Minimum Bounded Rectangle - MBR*) dando lugar ao uso de uma estrutura de indexação do tipo R-tree (Guttman, 1984). Mas, da mesma forma que no caso de cores, apresentam-se dificuldades com a invariância e a dimensionalidade do vetor de características, propiciando a “maldição da alta dimensionalidade” (Hinneburg et al., 1999).

2.3.3 Forma

Deve-se indicar que em sistemas de recuperação de imagens por conteúdo, o atributo forma é a abordagem que mostra maior dificuldade, especialmente pelo fato de ter que segmentar e conhecer o tamanho dos objetos contidos na imagem. Por exemplo, em aplicações médicas, a forma e tamanho dos tumores é de muita importância na hora de classificá-los como malignos ou benignos. Tumores com bordas irregulares têm uma alta probabilidade de serem malignos e aqueles que mostram bordas regulares geralmente são benignos (Theodoridis, 1999).

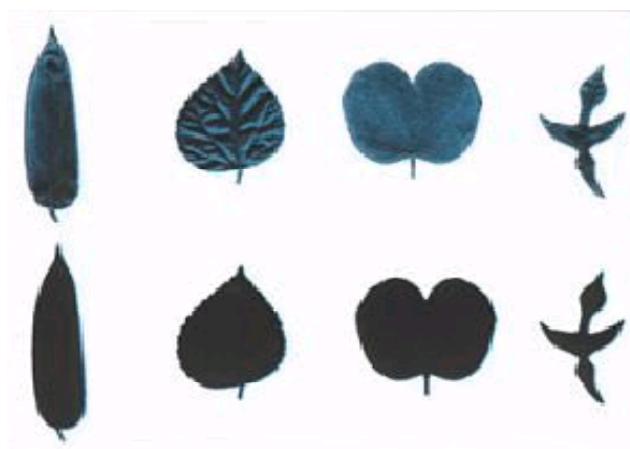


Figura 2.6: Formas de imagens (da Fontoura Costa and Cesar, 2000)

Vários esquemas foram propostos na literatura para a recuperação de imagens baseadas na forma. Uma técnica hierárquica que mistura métodos de momentos invariantes (Gonzales and Woods, 1993) com arco poligonal 2D é apresentado em (You and Bhattacharya, 2000). O sistema IRUS (Adoram and Lew, 1999) utiliza contornos ativos para a segmentação de objetos.

Uma análise multiescala de objetos bidimensionais aplicando Gabor *wavelets* é apresentado em (Cesar, 1997). Para um apanhado de técnicas de análise e classificação de formas recomenda-se (da Fontoura Costa and Cesar, 2000), onde o processo de extração de caracterís-

ticas de formas é definido com base em três tarefas: (a) organizar e visualizar as características; (b) extraír características; e (c) mensurar as características extraídas dos objetos. Na figura 2.6 apresentam-se alguns exemplos de forma como resultado do trabalho de (da Fontoura Costa and Cesar, 2000). Outro trabalho que aplica *wavelets* tentando melhorar o problema de variância na translação utiliza o que se denomina como momento máximo de *wavelets* (Do et al., 1999) demonstrando que também são invariantes à escala e rotação.

Existem duas abordagens pelas quais se pode obter a caracterização da imagem pela forma (Theodoridis, 1999). Uma é desenvolver técnicas que outorguem uma descrição total da borda do objeto (coeficientes de Fourier) e a outra é utilizar qualidades que descrevam as características morfológicas da região (número de esquinas na borda). Nessa abordagem estão as características baseadas em momentos. O trabalho recente de (Chávez, 2002) envolve esta última abordagem.

2.3.4 Resumo de Características de Baixo Nível

Na tabela 2.1 são brevemente descritos os tipos de características de baixo nível da imagem que têm sido amplamente usadas na recuperação de imagens por conteúdo, considerando-se suas vantagens e limitações. Em termos de cor, as características mais utilizadas são as dos histogramas de cor (Swain and Ballard, 1991) (Bueno, 2002), momentos de cor e vetores de coerência da cor. Essas características descrevem as propriedades globais de uma imagem e podem ser facilmente extraídas. Uma grande limitação é a sua impossibilidade de representar adequadamente a informação de localização espacial ou de objetos na imagem. As características morfológicas suplantam essa deficiência e descrevem melhor os objetos da imagem. Pode-se descrever a forma e as características de contorno de objetos basicamente de três maneiras: por aproximação polinomial, por momentos invariantes e por descritores de Fourier. As características morfológicas outorgam um alto nível de abstração em termos do formato dos objetos em uma imagem, mas isso requer o uso de bons algoritmos de segmentação para extraír objetos de interesse de uma imagem. As características de textura fornecem um nível intermediário de abstração em uma imagem, assim como as características da cor.

Na tabela 2.2 resumem-se alguns trabalhos anteriores em termos das diferentes características usadas nestes sistemas. Os espaços em branco da tabela representam informações não disponibilizadas. Alguns destes sistemas permitem a mistura de buscas textuais com buscas automáticas de imagens por conteúdo. Excetuando o Netra e o VisualSEEK, a maioria destes sistemas não têm suporte para indexação e consulta baseados na informação espacial (Albuz et al., 2001).

Atributo da Imagem	Características da imagem	Vantagens	Limitações
Cor	Histogramas, momentos e vetores de coerência	Podem ser facilmente extraídos automaticamente. Computacionalmente menos custosos. Representam as propriedades globais da imagem.	Não têm a possibilidade de representar a informação de localização espacial. Descrição limitada da semântica da imagem.
Textura	Matrizes de co-ocorrência, características de multiescala e características de filtros de Gabor.	Podem ser automaticamente extraídos da imagem. Outorgam informação global assim como local.	Podem ser computacionalmente caros (extração e casamento de características) e difíceis na definição. Descrição limitada da semântica da imagem.
Forma	Aproximação poligonal, momentos invariantes e descritores de Fourier.	Consegue obter um alto nível de abstração em termos morfológicos dos objetos locais. Permitem consultas no nível de objetos.	Não podem ser automaticamente extraídos (precisam de bons algoritmos de segmentação para extrair os objetos de interesse). Computacionalmente caros para lograr invariância às mudanças do objeto rígido.

Tabela 2.1: Descrição de tipos de características de baixo nível (Vailaya, 2000)

2.4 Exemplos de Características Extraídas de Imagens 2D

Em alguns casos práticos pode ser importante aplicar operações de pré-processamento com o objetivo de melhorar a qualidade da imagem, como por exemplo, remover ou atenuar o ruído nas imagens. Existem muitos trabalhos e métodos já desenvolvidos para esse fim, mas neste trabalho assume-se que as imagens já estarão pré-processadas e prontas para extrair as características.

A seguir serão apresentadas algumas caracterizações estatísticas que descrevem a forma do histograma, para em seguida detalhar aquelas que descrevem as regiões (textura e forma).

Sistema	Forma	Cor local	Cor global	Textura	Localização espacial
QBIC	✓	✗	✓	✓	✗
VisualSEEK	✓	✓	✓		✓
Virage	✓	✓	✓	✓	✗
PhotoBook	✓	✗	✗	✓	✗
Chabot	✓			✓	
MARS	✗		✓	✓	✗
Fast Multiresolução				✓	
Netra		✓		✓	✓
WBIIS	✗	✗		✓	

Tabela 2.2: Comparação de sistemas existentes (Albuz et al., 2001)

2.4.1 Estatística Descritiva para Imagens

As características mais básicas de toda imagem são as medidas que definem a amplitude da imagem em termos da luminância, valor espectral e outras unidades. A forma em que se apresenta o histograma de cor fornece informação sobre a natureza da imagem. Por exemplo, um histograma com uma distribuição estreita indicaria que se trata de uma imagem de contraste baixo. Um histograma bi-modal às vezes indica que a imagem contém um objeto que tem uma amplitude mais estreita em relação ao fundo da imagem. Essas medidas podem ser quantizadas como se descreve a seguir (Theodoridis, 1999).

Seja I a variável randômica que representa os níveis de cinza na região de interesse, o histograma $P(I)$ é definido como:

$$P(I) = \frac{\text{número de pixels com nível de cinza } I}{\text{número total de pixels na região}} \quad (2.1)$$

Isto significa que $P(I)$ é a fração de pixels com nível de cinza I . O gráfico de um histograma (onde $P(I)$ representa o eixo da porcentagem e I para $I = 0, 1, \dots, N_g - 1$ representa o eixo da intensidade) provê uma percepção da distribuição de cores.

As características que descrevem a forma de um histograma são muito úteis e são fornecidos pelos seguintes descritores estatísticos, entre outros (Loew, 2000) (Theodoridis, 1999):

a. Momentos

$$m_i = E[I^i] = \sum_{I=0}^{N_g-1} I^i P(I), \quad i = 1, 2, \dots \quad (2.2)$$

Onde $m_0 = 1$ representa o momento zero. Se $i = 1$ então $m_1 = E[I]$ que representa a **media** do valor I .

b. Momentos Centrais

$$\mu_i = E[(I - E[I])^i] = \sum_{I=0}^{N_g-1} P(I)(I - m_1)^i \quad (2.3)$$

Onde as variáveis envolvidas possuem o mesmo significado anterior. O segundo momento central μ_2 é a variância σ^2 é útil na descrição da uniformidade de uma determinada região da imagem.

O terceiro momento central é uma medida de não centralismo (*skewness*) que calcula o grau da assimetria do histograma ao redor da média, as vezes denotada como σ^3 . O quarto

momento central (*kurtosis*) mede a monotonia relativa da imagem. Dependendo do valor deste momento o histograma será denominado de *platykurtic*, para valores maiores, ou de *leptokurtic* para valores pequenos. No caso de uma distribuição normal é denominada como *mesokurtic*.

c. Energia

$$E = \sum_{I=0}^{N_g-1} [P(I)]^2 \quad (2.4)$$

d. Entropia

$$H = - \sum_{I=0}^{N_g-1} P(I) \log_2(I) \quad (2.5)$$

2.4.2 Descrição de Regiões

As regiões são tipicamente definidas com base na sua homogeneidade interna e em algumas características adicionais. Nesta situação a escala é importante na definição dessa homogeneidade. Devido à existência de auto similaridade na homogeneidade, características fractais podem outorgar informação que outros tipos de características não podem.

2.4.2.1 Textura e a matriz de co-ocorrência

As características resultantes dos descritores estatísticos mostrados em 2.4.1 fornecem de informação relacionada à distribuição dos níveis de cinza dentro da imagem, mas não outorgam informação acerca da posição relativa dos níveis de cinza distintos dentro da imagem. Estão todos os valores baixos do nível de cinza localizados juntos? ou, estão alternados com os valores altos? Esse tipo de informação pode ser extraído dos histogramas de segunda ordem, onde os pixels são considerados em pares. Dois parâmetros adicionais são considerados, aquele que determina a distância relativa entre os pixels e sua relativa orientação.

Seja d a distância relativa medida em número de pixels ($d = 1$ para pixels vizinhos, etc.). A orientação ϕ está quantizada em quatro direções: horizontal, diagonal, vertical e anti-diagonal ($0^\circ, 45^\circ, 90^\circ, 135^\circ$), como é indicado na figura 2.7. Para cada combinação de d e ϕ é definido um histograma de duas dimensões

$$0^\circ : P(I(m, n) = I_1, I(m \pm d, n) = I_2) \quad (2.6)$$

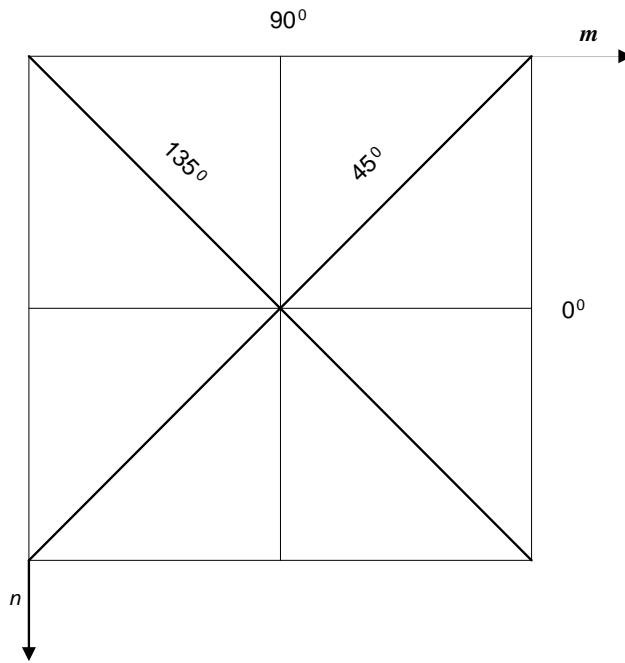


Figura 2.7: As quatro orientações usadas para construir as matrizes de co-ocorrência (Theodoridis, 1999)

$$= \frac{\text{número de pares de pixels na distância } d \text{ com valores } (I_1, I_2)}{\text{número total de pares possíveis}}$$

Em forma similar:

$$45^\circ : P(I(m, n) = I_1, I(m \pm d, n \mp d) = I_2)$$

$$90^\circ : P(I(m, n) = I_1, I(m, n \mp d) = I_2)$$

$$135^\circ : P(I(m, n) = I_1, I(m \pm d, n \pm d) = I_2)$$

Para cada um destes histogramas é definido um arranjo, conhecido como *matriz de co-ocorrência* ou *matriz de dependência espacial*. Seja, por exemplo, o arranjo da imagem $I(m, n)$

$$I = \begin{bmatrix} 0 & 0 & 2 & 2 \\ 1 & 1 & 0 & 0 \\ 3 & 2 & 3 & 3 \\ 3 & 2 & 2 & 2 \end{bmatrix} \quad (2.7)$$

A qual corresponde a uma imagem de 4×4 pixels. Também assuma-se que $N_g = 4$ ($I(m, n) \in \{0, 1, 2, 3\}$). A matriz de co-ocorrência para um par (d, ϕ) está definida como a matriz de $N_g \times N_g$

$$A = \frac{1}{R} \begin{bmatrix} \eta(0, 0) & \eta(0, 1) & \eta(0, 2) & \eta(0, 3) \\ \eta(1, 0) & \eta(1, 1) & \eta(1, 2) & \eta(1, 3) \\ \eta(2, 0) & \eta(2, 1) & \eta(2, 2) & \eta(2, 3) \\ \eta(3, 0) & \eta(3, 1) & \eta(3, 2) & \eta(3, 3) \end{bmatrix}$$

Onde $\eta(I_1, I_2)$ é o número de pixels pares, na posição relativa (d, ϕ) , a qual tem valores do nível de cinza I_1, I_2 , respectivamente. R é o número total de pixels pares. Assim $\frac{1}{R}\eta(I_1, I_2) = P(I_1, I_2)$. Para a imagem 2.7 e a posição relativa do pixel $(1, 0^\circ)$ tem-se:

$$A^0(d = 1) = \frac{1}{24} \begin{bmatrix} 4 & 1 & 1 & 0 \\ 1 & 2 & 0 & 0 \\ 1 & 0 & 6 & 3 \\ 0 & 0 & 3 & 2 \end{bmatrix}$$

O que quer dizer, que para cada intensidade par, como o par $(0, 0)$, conta-se o número de pixels pares na distância relativa $d = 1$ e orientação $\phi = 0^\circ$ que tomam esses valores. Para o exemplo apresentado esse valor é 4. Dois dos quais resultam da procura na direção positiva e os outros dois na direção negativa. Considerando a definição 2.6, estes pares de pixels têm as coordenadas (m, n) e $(m \pm 1, n)$ e os níveis de cinza $I_1 = 0, I_2 = 0$. O número total de pixels pares para este caso é de 24. De fato, para cada fila têm-se $N_x - 1$ pares e têm-se N_y filas. Assim, o número total nas duas direções (positivo e negativo) têm-se $2(N_x - 1)N_y = 2(3 \times 4) = 24$. Para a direção diagonal de 45° e $d = 1$ para cada linha têm-se $2(N_x - 1)$ pares, exceto para a primeira, que não tem nenhum par. Assim, o número total é $2(N_x - 1)(N_y - 1) = 2(3 \times 3) = 18$. Para $d = 1$ e 90° têm-se $2(N_y - 1)N_x$ pares e finalmente para $d = 1$ e 135° $2(N_x - 1)(N_y - 1)$. Para a imagem exemplo apresentada e $(d = 1, \phi = 45^\circ)$ obtém-se:

$$A^{45}(d = 1) = \frac{1}{18} \begin{bmatrix} 0 & 1 & 2 & 1 \\ 1 & 0 & 1 & 1 \\ 2 & 1 & 0 & 3 \\ 1 & 1 & 3 & 0 \end{bmatrix}$$

Apesar de o exemplo acima mostrar matrizes simétricas, é importante ressaltar que este não é o caso geral.

Tendo definido as probabilidades de ocorrência dos níveis de cinza com relação a posição espacial relativa do pixel, a seguir, estabelecem-se as características correspondentes a serem

extraídas. Algumas das características têm uma interpretação física direta em relação à textura, por exemplo, onde é possível quantificar a rugosidade, suavidade, etc. Por outro lado, outras características não têm tais propriedades, mas ainda assim codificam informação relacionada à textura com uma grande capacidade discriminatória (Theodoridis, 1999)(Loew, 2000).

Após a obtenção da matriz de co-ocorrência, as seguintes características podem ser calculadas:

- ***Segundo momento angular (SMA)***

$$SMA = \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} (P(i, j))^2 \quad (2.8)$$

Esta característica é uma medida da suavidade da imagem. De fato, se todos os pixels são do mesmo nível de cinza $I = k$, então $P(k, k) = 1$ e $P(i, j) = 0$, $i \neq k$ ou $j \neq k$ e $SMA = 1$. No outro extremo, tem-se todos os possíveis pares de níveis de cinza com a mesma probabilidade $\frac{1}{R}$, então $SMA = \frac{R}{R^2} = \frac{1}{R}$. Quanto menos suave seja a região, existirá maior distribuição uniforme $P(i, j)$ e menor o SMA .

- ***Contraste (CON)***

$$CON = \sum_{n=0}^{N_g-1} n^2 \left\{ \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} P(i, j) \right\}_{|i-j|=n} \quad (2.9)$$

Esta é a medida do contraste da imagem, isto é, a medida das variações locais dos níveis de cinza. De fato, $\sum_i \sum_j P(i, j)$ é a porcentagem de pares de pixels cujas intensidades se diferenciam por n . A dependência n^2 determina com maior valor as grandes diferenças. CON toma valores altos para imagens de alto contraste.

- ***Momento de diferenças inversas (IDF)***

$$IDF = \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} \frac{P(i, j)}{1 + (i - j)^2} \quad (2.10)$$

Esta característica toma valores altos para imagens de baixo contraste devido à dependência inversa $(i - j)^2$.

- ***Entropia (H)***

$$H_{xy} = - \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} P(i, j) \log_2 P(i, j) \quad (2.11)$$

A entropia é uma medida randômica e toma valores baixos para imagens suaves.

2.4.2.2 Forma

A representação e análise morfológica cumpre um papel central em diversas aplicações de visão computacional. De fato, procedimentos de reconhecimento ótico de caracteres, casamento de contornos para reconstrução tridimensional de imagens biomédicas, inspeção visual e muitas outras tarefas podem ser realizadas por processos baseados na forma. De maneira geral, existem duas abordagens básicas para a representação e a análise morfológica (Sonka et al., 1999): a abordagem por contornos e a por regiões. Exemplos de representação de formas baseadas em regiões são os descritores simples de região escalar, momentos, *convex hull*, os esqueletos, eixos de simetria e a decomposição morfológica. Por outro lado a aproximação poligonal, o código de cadeia, as primitivas geométricas, as curvas paramétricas, a representação B-spline, os descritores de Fourier e a transformada de Hough são exemplos de representações de formas baseadas em contornos (Cesar, 1997) (Sonka et al., 1999).

Primitivas geométricas

- Diâmetro efetivo (diâmetro da circunferência que tem a mesma área)

$$d = 2\sqrt{\frac{A}{\pi}} \quad (2.12)$$

Onde A é a área do objeto

- Circularidade ($C = 1$ para um círculo)

$$C = \frac{4\pi A}{P^2} \quad (2.13)$$

onde P é o perímetro.

- Solidez (máximo para um círculo)

$$solidez = \frac{P^2}{A} \quad (2.14)$$

- Projeção

Apesar de sua maior utilização no processamento de imagens binárias, as projeções podem servir como uma base para definir os descritores de regiões relacionadas. Estes descritores podem ser definidos em todas as direções, onde os mais usados são o horizontal e o vertical:

$$\text{Projeção horizontal} = p_h(i) = \sum_{j=1}^{N_g} I(i, j) \quad (2.15)$$

$$\text{Projeção vertical} = p_v(j) = \sum_{i=1}^{N_g} I(i, j) \quad (2.16)$$

Tais projeções também podem ser úteis para medir a homogeneidade em imagens de tons de cinza e a altura e largura das regiões.

Momentos

Quando duas imagens são comparadas entre si, ou com uma imagem padrão², um conjunto de momentos derivados por Hu (Hu, 1962) podem ser úteis. Estes momentos têm a característica de serem invariantes à translação, rotação e mudanças de escala. Para a imagem $I(x, y)$, define-se

$$\mu_{pq} = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (x_i - \bar{x})^p (y_j - \bar{y})^q I(x, y) \quad (2.17)$$

onde

$$\bar{x} = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} x_i I(x_i, y_j) \quad (2.18)$$

$$\bar{y} = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} y_j I(x_i, y_j) \quad (2.19)$$

e os momentos centrais normalizados são definidos como:

$$\eta_{pq} = \frac{\mu_{pq}}{(\mu_{00})^\gamma} \quad (2.20)$$

onde

$$\gamma = \frac{p+q}{2} + 1 \quad (2.21)$$

²As comparações também podem ser feitas entre regiões das imagens

Os momentos invariantes estão então definidos como:

$$\begin{aligned}\phi_1 &= \eta_{20} + \eta_{02} \\ \phi_2 &= (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \\ \phi_3 &= (\eta_{30} - 3\eta_{12})^2 + (\eta_{03} - 3\eta_{21})^2 \\ \phi_4 &= (\eta_{30} + \eta_{12})^2 + (\eta_{03} - \eta_{21})^2 \\ \phi_5 &= (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + \\ &\quad (\eta_{03} - 3\eta_{21})(\eta_{21} + \eta_{03})[(\eta_{21} + \eta_{03})^2 - 3(\eta_{30} + \eta_{12})^2] \\ \phi_6 &= (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}) \\ \phi_7 &= (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + \\ &\quad (\eta_{30} - 3\eta_{12})(\eta_{21} + \eta_{03})[(\eta_{21} + \eta_{03})^2 - 3(\eta_{30} + \eta_{12})^2]\end{aligned}$$

Outros tipos de momentos são os de Zernike que estão baseados em funções polinomiais complexas, formando um *conjunto ortonormal completo sobre o círculo unitário* $x^2 + y^2 \leq 1$ e são definidos como:

$$V_{pq}(x, y) = V_{pq}(\rho, \theta) = R_{pq}(\rho) \exp(jq\theta) \quad (2.22)$$

Onde:

p é um inteiro não negativo.

q é um inteiro sujeito à restrição $p - |q|$ par, $|q| \leq p$

$$\rho = \sqrt{x^2 + y^2}$$

$$\theta = \operatorname{tg}^{-1} \frac{y}{x}$$

$$R_{pq}(\rho) = \sum_{s=0}^{(p-|q|)/2} \frac{(-1)^s [(p-s)!] \rho^{p-2s}}{s! (\frac{p+|q|}{2}-s)! (\frac{p-|q|}{2}-s)!}$$

Os momentos de Zernike de uma função $I(x, y)$ estão dados por:

$$A_{pq} = \frac{p+1}{\pi} \int \int_{x^2+y^2 \leq 1} I(x, y) V_{pq}^*(\rho, \theta) dxdy \quad (2.23)$$

onde $*$ denota a conjugação complexa. Para uma imagem digital, os respectivos momentos de Zernike são calculados como:

$$A_{pq} = \frac{p+1}{\pi} \sum_i I(x_i, y_i) V_{pq}^*(\rho_i, \theta_i), \quad x_i^2 + y_i^2 \leq 1 \quad (2.24)$$

Onde i toma todos os pixels da imagem. O cálculo dos momentos correspondentes de uma imagem, considera o centro da imagem como a origem e os pixels são mapeados dentro do

círculo unitário, isto é, $x_i^2 + y_i^2 \leq 1$. Os pixels que não se encontram dentro do círculo unitário não são considerados. A magnitude dos momentos de Zernike é invariante à rotação.

Como o presente trabalho visa a utilização de *wavelets* para a extração de características de imagens, é importante colocar que um dos maiores problemas da transformada *wavelet* é a dificuldade de tratar imagens que apresentam mudanças de translação, rotação e escala. No entanto, Minh Do (Do et al., 1999) utiliza uma abordagem interessante que aplica o módulo máximo da transformada *wavelet* para contornar esse problema. Para medir a similaridade entre representações *wavelet* máxima, aplica-se a diferença de momentos.

2.5 Considerações Finais

Nesse capítulo foram mostradas algumas das principais técnicas de extração de características de imagens baseadas em seus atributos principais (cor, textura e forma). As imagens médicas geralmente são apresentadas em tons de cinza, pelo que, para o presente trabalho foram considerados, especialmente, as características de cor e textura.

O maior problema no que se refere ao processo de extração de características é a procura daquele espaço de características mais representativo do domínio de imagens, para assim permitir a recuperação de imagens de maneira mais eficiente. Este é um assunto em aberto onde muita pesquisa tem sido feita nos últimos anos (Jain and Vailaya, 1998).

A seguir, no próximo capítulo, apresentam-se a base teórica e a maneira de ser aplicada a técnica sobre a qual está baseado este trabalho, a transformada *wavelet*. Essa abordagem foi utilizada para transformar as imagens em espaços de freqüência multiescalas, sobre os quais aplicou-se medidas estatísticas para gerar os vetores de características.

Introdução à Teoria de *Wavelets*

3.1 Considerações Iniciais

Neste capítulo apresenta-se uma breve introdução à teoria de *wavelets* sob a ótica da análise de sinais. A todo momento estamos nos deparando com sinais, por exemplo: sinais sísmicos, sinais de rádio, áudio, fala humana, dados financeiros, imagens médicas entre outros.

Esses sinais, para seu processamento e análise, contam com uma série de ferramentas, onde talvez a mais conhecida seja a transformada de Fourier, que representa um sinal através de funções senoidais de diferentes freqüências. Através da transformada de Fourier, um sinal representado no espaço do tempo é transformado no espaço de freqüência.

A transformada *wavelet* apresenta-se como uma ferramenta alternativa para o processamento de sinais, mudando o paradigma de representação dos mesmos, ao utilizar funções base de *wavelets* (em vez das senoidais de Fourier) para transformar um sinal no espaço de tempo-escala.

No caso de processamento de imagens, um dos pontos fortes da transformada *wavelet* vem do fato que com uma quantidade mínima de valores (em uma escala menor da imagem) é possível representar a imagem toda. Esses valores podem ser aproveitados na montagem do vetor de características para representação de uma imagem.

As *wavelets* estão sendo utilizadas em inúmeras áreas de aplicação, por exemplo na mecânica de fluidos, computação numérica, análise de imagens, processamento de sinais, sistemas de controle, fenômenos biológicos, medicina, medidas e psicologia. No futuro próximo espera-se

que a análise de *wavelets* seja uma ferramenta de análise e visualização como a transformada de Fourier. A seguir apresenta-se uma breve introdução a esta interessante teoria, onde a abordagem utilizada é de multiresolução e no final do capítulo é desenvolvida a teoria dos filtros de Gabor, os quais são aproveitados para caracterizar textura.

3.2 Introdução Histórica

As *wavelets* são conhecidas dos matemáticos há muito tempo, mas a formalização desta teoria foi realizada recentemente (na década de 80) (Grossmann and Morlet, 1984) (Mallat, 1989) (Daubechies, 1990). O livro de Yves Meyer (Meyer, 1993) aborda os aspectos históricos das *wavelets*, desde Fourier (1807) até as contribuições recentes de A. Grossmann, J. Morlet, S. Mallat e outros. Outro trabalho muito recomendado é de Barbara B. Hubbard (Hubbard, 1998) que descreve o nascimento, a história e os conceitos originários das *wavelets* de uma forma clara.

A primeira menção que poderia-se chamar de *wavelet* é aquela que foi apresentada no anexo da tese de Alfred Haar em 1909 e que agora é chamada de base de Haar. O termo “*wavelets*” é atribuído a Norman Ricker (1940) no seu trabalho de sismologia, mas foi introduzido por J. Morlet, sendo sua base matemática formalizada pelo físico teórico A. Grossmann (Grossmann and Morlet, 1984). Os dados sísmicos estudados por Morlet exibiam conteúdos de freqüência que mudavam rapidamente ao longo do tempo, além disso mostraram que qualquer tipo de sinal pode ser analisado em termos de escalas e translações de uma simples função *wavelet mãe*. Yves Meyer (Meyer, 1993) e Stephane Mallat (Mallat, 1988) desenvolveram esta idéia em uma teoria denominada por *análise de multiresolução*. Em 1989 Mallat (Mallat, 1989) mostrou que a análise de multiresolução pode ser vista simplesmente como uma forma de *algoritmos de pirâmide* usados no processamento de imagens.

3.3 Transformada de Fourier

Como foi especificado no início do capítulo, uma das técnicas mais populares no processamento de sinais é a transformada de Fourier, que tem como objetivo transformar um sinal (função) do domínio de espaço para o domínio de freqüência (Wang, 2001), (Gonzales and Woods, 1993).

A função responsável pela transformação é dada por

$$F[u] = \int f[t]e^{-i2\pi ut}dt \quad (3.1)$$

Esta equação corresponde à transformada de Fourier de um sinal contínuo $f(t)$, a qual apresenta algumas deficiências, entre elas o fato de que esta transformação não está localizada no domínio do espaço, então ela não pode representar adequadamente as mudanças que acontecem no sinal no espaço. Isto é devido ao fato de que a transformada está baseada na integração de toda a função para o cálculo de cada freqüência. Esse não seria um problema se o sinal não mudasse durante o tempo (sinais estacionários) mas acontece que muitos sinais interessantes contém muitas características não estacionárias ou transitórias (fluxos, tendências, mudanças repentinas, início e final de um evento). Essas características geralmente são a parte mais importante do sinal.

Para solucionar esse problema Dennis Gabor (1946) adaptou a transformada de Fourier para analisar só uma pequena porção do sinal em um tempo, introduzindo a chamada transformada por janelas de Fourier (*Windowed Fourier Transform - WFT*). Neste caso, uma janela de observação é deslocada no domínio do tempo e a transformada de Fourier da porção “visível” do sinal é calculada para cada posição da janela. Considerando $j(t)$ a função que cumpre o papel de janela, pode-se definir a transformada por janelas de Fourier de um sinal contínuo $f(t)$ como:

$$F(u, b) = \int j(t - b)f[t]e^{-i2\pi ut} dt \quad (3.2)$$

O problema é que uma vez que foi definido um tamanho para a janela de tempo, esta janela permanece constante para todas as freqüências. Pode acontecer que alguns sinais precisem de um enfoque mais flexível, onde o tamanho da janela possa mudar com o objetivo de detectar o conteúdo local da freqüência (Daubechies, 1992) (Gomez et al., 1997).

3.4 Transformada Wavelet

Depois da transformada por janelas de Fourier, a utilização das *wavelets* é o passo lógico seguinte. Logo, pode-se interpretá-la como uma técnica por janelas com regiões de dimensão variável, onde as *wavelets*, diferentemente de Fourier, têm como base uma função de duração limitada, isto é, de suporte compacto, que é uma propriedade na qual seu domínio é diferente de zero em uma extensão finita e igual a zero em todo o resto. Isto torna interessante a utilização das *wavelets* no caso específico da análise de imagens, pois as mudanças de regiões ou bordas podem ser detectadas mais facilmente.

A definição de uma transformada *wavelet* considerando um sinal contínuo é dada por:

$$F(a, b) = \int f(t)\psi_{a,b}(t)dt \quad (3.3)$$

Nesta equação os parâmetros a e b variam continuamente em \mathbb{R} , sendo que as funções $\psi_{a,b}$ são denominadas *wavelets* e definidas da seguinte forma:

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}}\psi\left(\frac{t-b}{a}\right) \quad (3.4)$$

A transformada *wavelet* para sinais discretos é definida como:

$$F_{m,n}(a, b) = a_0^{-m/2} \int f(t)\psi(a_0^{-m}t - nb_0) \quad (3.5)$$

Pode-se ver claramente que o comportamento desta função está baseada em dilatações e translações a partir de uma *wavelet mãe* ψ . Em ambos os casos, essa *wavelet mãe*, deve satisfazer a propriedade:

$$\int \psi(t)dt = 0 \quad (3.6)$$

Observando a equação 3.3, percebe-se que a transformada *wavelet* depende de dois parâmetros a e b , que correspondem às informações de escala e tempo respectivamente (Daubechies, 1992) (Gomez et al., 1997) (Misiti et al., 2001)

Obter os coeficientes de *wavelets* em cada escala possível requer uma grande quantidade de cálculo, tornando muito tedioso o trabalho (transformada contínua de *wavelets*). Devido a este fato é que a transformada discreta de *wavelets* escolhe um subconjunto de escalas e locações sobre os quais vai realizar os cálculos.

Segundo Daubechies (Daubechies, 1992) dentro da transformada discreta de *wavelets* distinguem-se duas abordagens: sistemas redundantes discretos (*frames*) e ortonormal (e outras) bases de *wavelets*. A segunda abordagem considera a estratégia de análise de multiresolução, desenvolvida por Mallat (Mallat, 1989), a qual será considerada como base do presente trabalho.

3.4.1 Métodos Multiresolução

As primeiras construções de bases ortonormais de *wavelets* pareciam um pouco miraculosas. A situação mudou com a chegada da análise de multiresolução, formulada em 1986 por Mallat e Meyer. A multiresolução fornece um referencial onde bases de *wavelets* são naturalmente compreendidas, bem como permite a construção de novas bases. Quando Mallat trabalhou com *wavelets* de Meyer pela primeira vez, ele estava trabalhando com análise de imagens, onde a idéia de se estudar imagens em várias escalas simultaneamente era popular. Isto o estimulou a ver bases ortonormais de *wavelets* como uma ferramenta para descrever matematicamente o

“incremento na informação” necessário para se ir de uma aproximação grosseira para uma aproximação com maior resolução (Daubechies, 1992).

Uma decomposição multiresolução permite obter uma interpretação invariante de escala da imagem. A escala de uma imagem muda em razão da distância entre a cena e o centro ótico da câmera. Quando a escala da imagem é modificada, a interpretação da imagem não deveria mudar. Uma representação multiresolução pode ser parcialmente invariante de escala se a seqüência de parâmetros da resolução $(V_j)_{j \in \mathbb{Z}}$ varia exponencialmente (Mallat, 1989).

Para um melhor entendimento, utiliza-se o conceito de “espaço vetorial” da álgebra linear (ver apêndice A). Um *espaço vetorial* V é basicamente uma coleção de “objetos” (nomeados *vetores*, neste contexto) para os quais a adição e o produto escalar estão definidos. Assim, é possível adicionar dois vetores, escalar um vetor por alguma constante e assim sucessivamente (Stollnitz et al., 1996).

Agora é preciso definir uma base para cada espaço vetorial V_j ¹. As funções base para o espaço V_j são denominadas *funções escalares* e são usualmente denotadas pelo símbolo ϕ . Uma base simples para V_j é dada pelo conjunto de *funções caixa* escaladas e transladadas:

$$\phi_i^j(x) := \phi(2^j x - i) \quad i = 0, \dots, 2^j - 1$$

onde

$$\phi(x) := \begin{cases} 1 & \text{se } 0 \leq x < 1 \\ 0 & \text{se caso contrário} \end{cases}$$

A figura 3.1 mostra duas aproximações de uma função contínua por funções caixa.

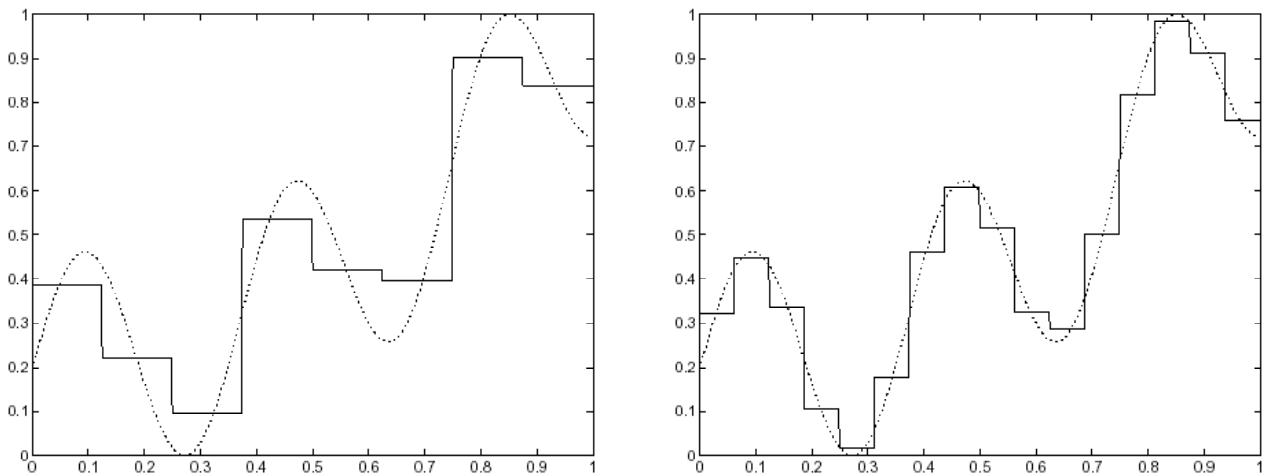


Figura 3.1: Uma função contínua $f(x)$ (linhas pontilhadas) e a função de aproximação caixa (linhas sólidas) em duas resoluções. Na esquerda a aproximação grosseira $A_0 f(x)$ e na direita é a aproximação de maior resolução $A_1 f(x)$ (Davis and Nosratinia, 1998).

¹Uma *base* para um espaço vetorial, em uma definição aproximada, consiste de um conjunto mínimo de vetores linearmente independentes dos quais todos os demais vetores no espaço vetorial podem ser gerados através de combinações lineares.

O passo seguinte na construção da análise de multiresolução consiste em escolher um *produto interno* definido sobre os espaços de vetores V_j . Para o exemplo mostrado, o produto interno “padrão” é dado por:

$$\langle f | g \rangle := \int_0^1 f(x) g(x) dx \quad (3.7)$$

Dois vetores u e v são ditos *ortogonais* sob um produto interno escolhido se $\langle u | v \rangle = 0$. Então, agora é possível definir um novo espaço vetorial W_j como o *complemento ortogonal* de V_j em V_{j+1} . Em outras palavras, W_j é o espaço de todas as funções em V_{j+1} que são ortogonais a todas as funções em V_j no produto interno escolhido.

Uma coleção de funções $\psi_i^j(x)$ que geram W_j são nomeadas **wavelets**. Estas funções base tem duas propriedades importantes:

- as funções base $\psi_i^j(x)$ de W_j , juntas às funções base ϕ_i^j de V_j , formam uma base para V_{j+1} ,
- cada função base $\psi_i^j(x)$ de W_j é ortogonal a cada função base ϕ_i^j de V_j no produto interno escolhido.

A figura 3.2 ilustra graficamente a construção de V_j e W_j .

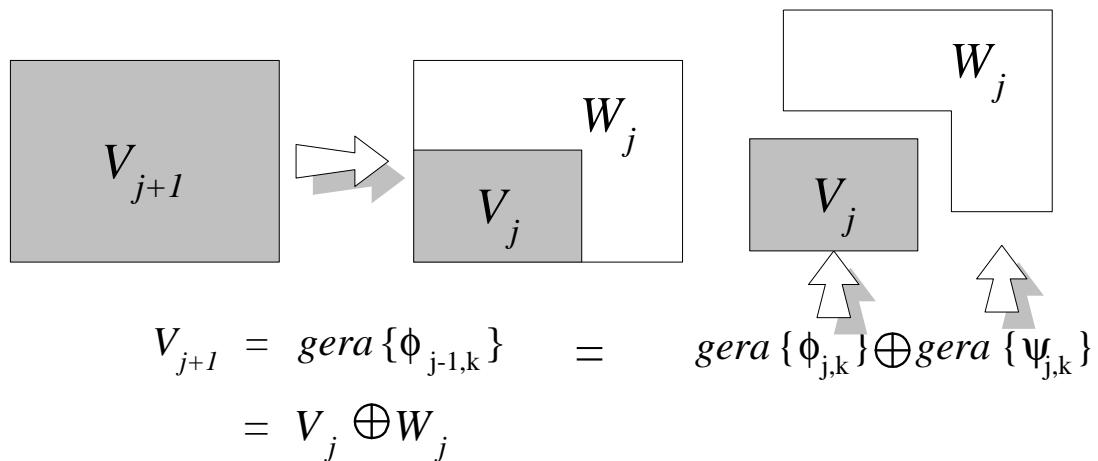


Figura 3.2: Geração do espaço superior em função de $\phi(x)$ e $\psi(x)$

Em resumo (Daubechies, 1992), a análise de multiresolução consiste de uma seqüência de espaços de aproximação sucessivos V_j . Mais precisamente, os subespaços fechados V_j satisfazem:

$$\dots V_{-2} \subset V_{-1} \subset V_0 \subset V_1 \subset V_2 \dots \quad (3.8)$$

com

$$\overline{\bigcup_{j \in \mathbb{Z}} V_j} = L^2(\mathbb{R}) \quad (3.9)$$

$$\bigcap_{j \in \mathbb{Z}} V_j = \{0\} \quad (3.10)$$

Todos os espaços são uma versão escalada do espaço central V_0 ,

$$f(x) \in V_j \Leftrightarrow f(2^j x) \in V_0 \quad (3.11)$$

e

$$f(x) \in V_0 \Rightarrow f(x - n) \in V_0, \text{ para todo } n \in \mathbb{Z} \quad (3.12)$$

indica a invariância de V_0 sob translações inteiras.

Agora precisa-se saber que $\phi \in V_0$

$$\phi_{0,n}(x) = \phi(x - n), n \in \mathbb{Z} \text{ constitui uma base ortonormal para } V_0 \quad (3.13)$$

A propriedade indicada pela equação 3.11 expressa que todos os espaços estão relacionados por escala a um mesmo espaço V_0 (este é o aspecto da “multiresolução”). Ainda por causa desta propriedade, se $f(x) \in V_j$, então, $f(x - 2^j n) \in V_j$, para todo $n \in \mathbb{Z}$. As condições indicadas pelas equações 3.13 e 3.11 implicam que $\{\phi_{j,n}\}_{j,n \in \mathbb{Z}}$ é uma base ortonormal para V_j para todo $j \in \mathbb{Z}$. Se definirmos P_j como o operador projeção ortogonal sobre V_j , a condição 3.10 assegura que $\lim_{j \rightarrow \infty} P_j f = f$ para todo $f \in L^2(\mathbb{R})$.

Cada V_j pode ser interpretado como um espaço de aproximação sucessiva: a aproximação de $f \in L^2(\mathbb{R})$ na resolução 2^j é definida como a projeção de f sobre V_j e quanto maior j , mais fina é a resolução obtida. A condição 3.11 significa que nenhuma escala é privilegiada. Os detalhes adicionais necessários para aumentar a resolução de 2^j para 2^{j+1} são dados pela projeção de f sobre o complemento ortogonal de V_j em relação a V_{j+1} o qual denotamos por W_j : $V_j \oplus W_j = V_{j+1}$ (Daubechies, 1992).

3.4.2 Transformada Wavelet de Haar

A análise com *wavelets* pode estar baseada no enfoque desenvolvido por Haar (Meyer, 1993). Em 1909, A. Haar descreveu uma base ortonormal de *wavelets* definida sobre o domínio $[0, 1]$, ou seja $h_0(x), h_1(x), \dots, h_n(x), \dots$, outras bases, além das de Fourier, tal que para qualquer função contínua $f(x)$ sobre o intervalo $[0, 1]$, a série

$$\sum_{j=1}^{\infty} \langle f, h_j \rangle h_j(x) \quad (3.14)$$

converge a $f(x)$ de forma uniforme sobre $[0, 1]$. Aqui, $\langle u, v \rangle$ denota o produto interno de u e v :

$$\langle u, v \rangle = \int_0^1 u(x) \overline{v(x)} dx \quad (3.15)$$

onde \overline{v} é o conjugado complexo de v o qual é igual a v se a função é real.

Uma versão da construção de Haar é a seguinte:

$$h(x) := \begin{cases} 1 & , \quad x \in [0, 0.5) \\ -1 & , \quad x \in [0.5, 1) \\ 0 & , \quad \text{caso contrario} \end{cases} \quad (3.16)$$

$$h_n(x) = 2^{j/2} h(2^j x - k) \quad (3.17)$$

onde $n = 2^j + k, k \in [0, 2^j], x \in [k2^{-j}, (k+1)2^{-1}]$.

Existem limitações no uso da construção de Haar. Como as funções base de Haar são funções de passos descontínuos, não são muito adequadas para a análise de funções estáveis com derivações contínuas. Visto que as imagens geralmente contêm regiões estáveis, a transformada *wavelets* de Haar não provê resultados satisfatórios em muitas aplicações com imagens (Wang, 2001).

3.4.3 Transformada Wavelet de Daubechies

Outro tipo de base para *wavelets* é a proposta por Daubechies (Daubechies, 1992). Para cada inteiro r , a base ortonormal para $L^2(\mathbb{R})$ está definida como

$$\phi_{r,j,k}(x) = 2^{j/2} \phi_r(2^j x - k), j, k \in \mathbb{Z} \quad (3.18)$$

onde a função $\phi_r(x)$ em $L^2(\mathbb{R})$ tem a propriedade que $\phi_r(x - k)|k \in \mathbb{Z}$ é uma base seqüencial ortonormal em $L^2(\mathbb{R})$. Aqui, j é o índice de escala, k é o índice de translação e r é o índice de filtragem.

Então a *tendência* f_j na escala 2^{-j} de uma função $f \in L^2(\mathbb{R})$ está definida como

$$f_j(x) = \sum_k \langle f, \phi_{r,j,k} \rangle \phi_{r,j,k}(x) \quad (3.19)$$

Os *detalhes* ou *flutuações* são definidos por

$$d_j(x) = f_{j+1}(x) - f_j(x) \quad (3.20)$$

Para analisar estes detalhes em uma dada escala, define-se uma base ortonormal $\psi_r(x)$ com propriedades similares às de $\phi_r(x)$ descritas anteriormente.

As funções $\phi_r(x)$ e $\psi_r(x)$, denominadas por *wavelet pai* e *wavelet mãe* respectivamente, são as funções *wavelets* protótipo necessárias para a análise de *wavelets*. A figura 3.3 mostra algumas *wavelets* mães e pais. As famílias de *wavelets*, como aquelas definidas na equação 3.18, são geradas a partir da *wavelet* pai ou da mãe mudando a escala e translação no tempo (ou espaço em processamento de imagens).

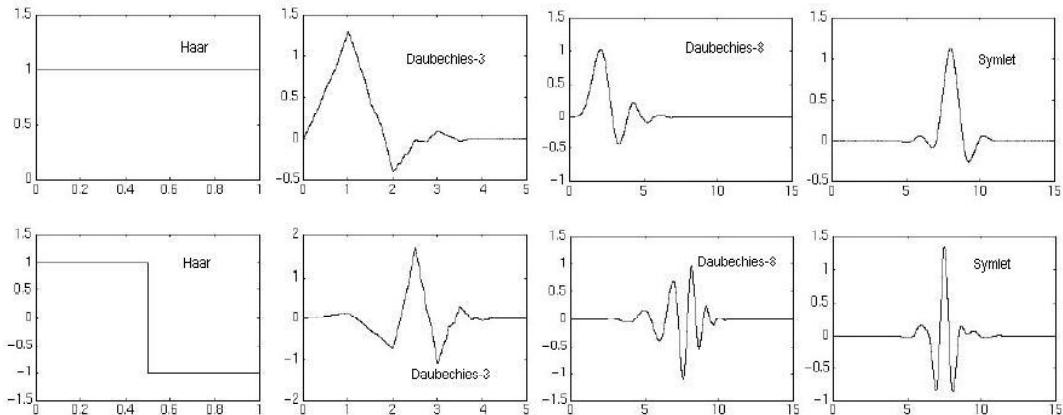


Figura 3.3: Algumas análise *wavelets*. Primeira fila: *wavelets* pai, $\phi(x)$. Segunda fila: *wavelets* mãe, $\psi(x)$ (Daubechies, 1992)

A base ortonormal de Daubechies tem as seguintes propriedades:

- ψ_r tem o intervalo de suporte compacto $[0, 2r + 1]$;
- ψ_r tem ao redor de $r/5$ derivações contínuas;
- $\int_{-\infty}^{\infty} \psi_r(x) dx = \dots = \int_{-\infty}^{\infty} x^r \psi_r(x) dx = 0$.

As *wavelets* de Daubechies obtém excelentes resultados no processamento de imagens devido às propriedades acima indicadas (Wang, 2001). A funções *wavelets* com suporte compacto podem ser implementadas mediante um filtro de longitude finita. Além disso, o suporte compacto permite a localização no domínio espacial. Devido ao fato que as funções base de *wavelets* têm derivações contínuas, é possível decompor uma função contínua mais eficientemente com os objetos da borda previstos. As *wavelets* mãe são utilizadas para determinar os detalhes de um sinal e a informação de tendência é armazenada nos coeficientes obtidos pelas *wavelets* pai.

As *wavelets* de Daubechies são usualmente implementadas numericamente através de filtros de espelho quadráticos (Meyer, 1993). A análise de multiresolução da tendência e flutuação de uma função é implementada mediante sua convolução com um filtro passa-baixa e um filtro passa-alta que são versões da mesma *wavelet*. A transformada *wavelets* de Haar é um caso especial da *wavelets* Daubechies com $r = 2$, onde as equações:

$$F_0(x(n)) = \frac{1}{\sqrt{2}}(x(n) + x(n+1)) \quad (3.21)$$

$$F_1(x(n)) = \frac{1}{\sqrt{2}}(x(n) - x(n+1)) \quad (3.22)$$

provêem a transformada do sinal $x(n)$, $n \in \mathbb{Z}$ mediante a *wavelet* Haar. Os filtros passa-baixa (equação 3.23) e passa-alta (equação 3.24) correspondentes são:

$$\left\{ \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right\} \quad (3.23)$$

$$\left\{ \frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}} \right\} \quad (3.24)$$

As transformadas *wavelet* de Daubechies com $r > 2$ apresentam uma melhor concentração da energia preservando mais a informação de tendência nos sinais ao se considerar apenas a parte do filtro passa-baixa. Por exemplo para $r = 4$ a transformada é denominada de *wavelet* de Daubechies-4 onde os filtros de passa baixa (equação 3.25) e de passa alta (equação 3.26) correspondentes são (Peters and Williams, 1998):

$$\left\{ \frac{(\sqrt{3}+1)}{4\sqrt{2}}, \frac{\sqrt{3}(\sqrt{3}+1)}{4\sqrt{2}}, \frac{(\sqrt{3}-1)}{4\sqrt{2}}, \frac{\sqrt{3}(\sqrt{3}-1)}{4\sqrt{2}} \right\} \quad (3.25)$$

$$\left\{ -\frac{\sqrt{3}(\sqrt{3}-1)}{4\sqrt{2}}, -\frac{(\sqrt{3}-1)}{4\sqrt{2}}, \frac{\sqrt{3}(\sqrt{3}+1)}{4\sqrt{2}}, -\frac{(\sqrt{3}+1)}{4\sqrt{2}} \right\} \quad (3.26)$$

3.4.4 Extensões para Dimensões Maiores

As transformadas *wavelet* uni-dimensionais descritas até agora podem ser extendidas para dimensões maiores de várias maneiras.

Uma forma bastante natural, de complexidade razoável, de extender a transformada para domínios 2D é através do produto tensorial (Gomez et al., 1997), (Stollnitz et al., 1996):

$$V_0 = v_0^h \otimes v_0^v$$

de forma que

$$F \in V_j \leftrightarrow F(2^j x_1, 2^j x_2) \in V_0$$

A função de escala ϕ_j é então definida como:

$$\phi_j(x_1, x_2) = 2^j \phi(2^j x_1 - k_1) \phi(2^j x_2 - k_2) \quad (3.27)$$

A relação de multiresolução precisa ser agora melhor analisada, já que é resultado de um produto tensorial:

$$\begin{aligned} V_{j+1} &= v_{j+1}^h \otimes v_{j+1}^v \\ &= v_j^h \oplus w_j^h \otimes v_j^v \oplus w_j^v \\ &= (v_j^h \otimes v_j^v) \oplus [(v_j^h \otimes w_j^v) \oplus (w_j^h \otimes v_j^v) \oplus (w_j^h \otimes w_j^v)] \end{aligned} \quad (3.28)$$

Observa-se então o aparecimento de três diferentes tipos de funções de *wavelets*: duas componentes cruzadas e uma componente pura. É possível interpretar como componentes *wavelets* horizontais (quando compostas com funções de escala verticais), *wavelets* verticais (quando compostas com funções de escala horizontais) e *wavelets* diagonais (quando são o produto tensorial de funções de *wavelets* horizontais e verticais) (Gomez et al., 1997):

$$\begin{aligned} \psi^h(x_1, x_2) &= \psi(x_1) \phi(x_2) \\ \psi^v(x_1, x_2) &= \phi(x_1) \psi(x_2) \\ \psi^d(x_1, x_2) &= \psi(x_1) \psi(x_2) \end{aligned} \quad (3.29)$$

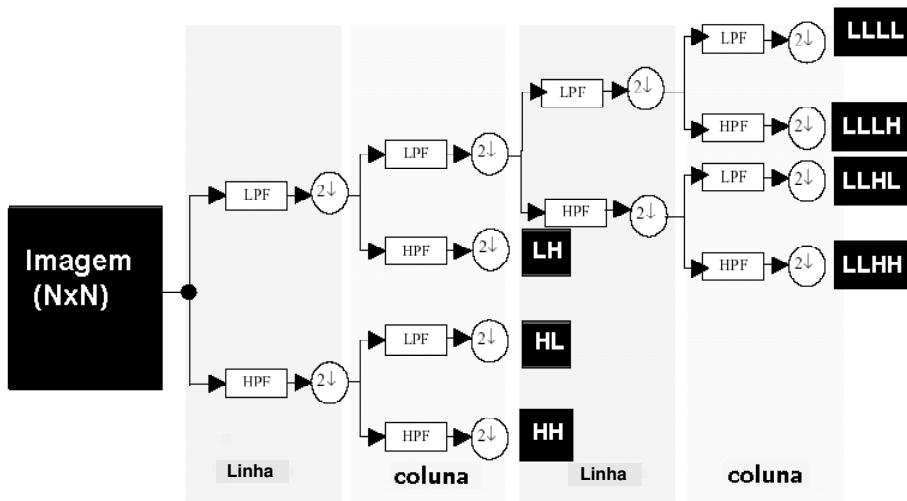


Figura 3.4: Processo de decomposição ao aplicar a transformada *wavelet* em um domínio 2D (Albuz et al., 2001).

O algoritmo para a transformada *wavelet* bi-dimensional explora a propriedade de separabilidade advinda da sua construção pelo produto tensorial. Dessa maneira, tanto o algoritmo de decomposição (figura 3.4 para 2D), quanto o algoritmo de reconstrução, podem ser implementados em duas dimensões com a aplicação dos algoritmos uni-dimensionais correspondentes, nas linhas e nas colunas da imagem. O resultado pode ser apreciado na imagem 3.5 após dois níveis de decomposição (Stollnitz et al., 1996).

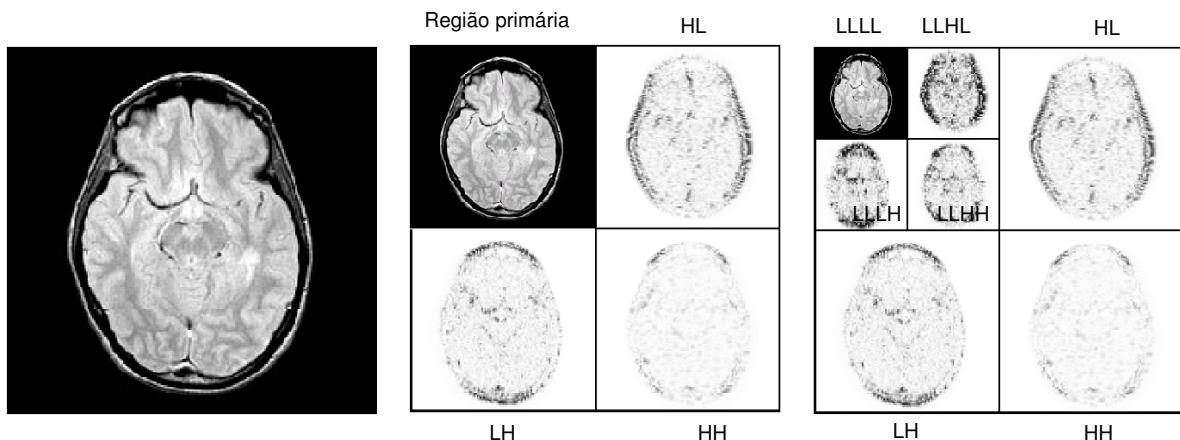


Figura 3.5: Transformada *wavelet* sobre um domínio 2D. Inicia-se com a imagem original e, após transformala, obter uma imagem com dois níveis de decomposição.

3.5 Transformada Wavelet de Gabor

Nesta seção utiliza-se do trabalho de Manjunath (Manjunath and Ma, 1996) o conceito de *transformada wavelet de Gabor*, o qual está baseado na ortogonalização dos filtros de Gabor e o aproveitamento da abordagem de multiresolução (Chang and Kuo, 1993) e (Mao and Jain, 1992). Esta abordagem é utilizada para caracterizar as imagens médicas que são motivo do presente trabalho.

A transformada *wavelet* de Gabor (*Gabor Wavelets Transform - GWT*) originalmente proposta como funções Gabor por (Gabor, 1946), tem conseguido resultados promissores quando utilizadas em aplicações de reconhecimento de textura e objetos (Tuceryan and Jain, 1993). As *wavelets* de Gabor são especialmente apropriadas para representação de características locais pelo fato de apresentar as seguintes propriedades (He et al., 2001): (a) As *wavelets* de Gabor são as melhores *wavelets* localizadas no tempo e na freqüência, (b) As *wavelets* de Gabor contêm um maior número de parâmetros, e (c) pesquisas em psicologia mostram que reações de células simples no córtex visual podem ser modeladas mediante as funções Gabor (Daugman, 1988) (Petkov and Kruizinga, 1997).

A utilização das *wavelets* de Gabor na extração de características está motivada por vários fatores. A representação de Gabor tem sido demonstrada como ótima no sentido de minimizar a incerteza das articulações bi-dimensionais no espaço e freqüência (Daugman, 1988). Estes filtros podem ser considerados tonalizáveis na orientação e na escala. Dessa forma, detectores de linhas e bordas, e cálculos estatísticos sobre estas micro-características em uma determinada região são usualmente efetuados para caracterizar a informação de textura subjacente (Manjunath and Ma, 1996), (Jain and Farrokhnia, 1991), (Bovik et al., 1990), (Turner, 1986).

As funções Gabor têm sido utilizadas no processamento de imagens em diversas formas, como na representação total da imagem mediante a definição de um conjunto de *wavelets* de Gabor 2D, que fornecem uma representação completa de uma imagem qualquer (Lee, 1996). Em (Manjunath and Ma, 1996) e (Jain and Healey, 1998) a média e o desvio-padrão dos coeficientes das GWT foram usados com sucesso na recuperação de imagens baseadas na textura. Utilizando esta técnica, pôde-se efetuar o reconhecimento de íris com grande sucesso, como é demonstrado no trabalho de Daugman (Daugman, 2002), assim como em impressões digitais (Jain et al., 2001),(Lee et al., 2001). Na área de reconhecimento de faces, Feris (Feris et al., 2001) desenvolveu uma técnica que analisa os subespaços *wavelets* gerados pela GWN (*Gabor Wavelets Network*) (Krueger and Sommer, 1999), para o reconhecimento de faces em seqüências de vídeo.

A seguir, apresentaremos a técnica de GWT para caracterização de imagens baseadas na textura, descrevendo suas principais características.

3.5.1 Funções Gabor e Wavelets

Uma função bi-dimensional de Gabor $\psi(x, y)$, usada como a *wavelet* mãe, é definida como (Manjunath and Ma, 1996):

$$\psi(x, y) = \left(\frac{1}{2\pi\sigma_x\sigma_y} \right) \exp \left[-\frac{1}{2} \left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2} \right) + 2\pi jWx \right] \quad (3.30)$$

onde σ_x e σ_y são os desvios-padrões de $\psi(x, y)$ ao longo dos eixos x e y , respectivamente. A constante W determina o comprimento de banda de freqüência dos filtros. As partes real e imaginária de $\psi(x, y)$ são mostradas na figura 3.6 (Rubner and Tomasi, 2000).

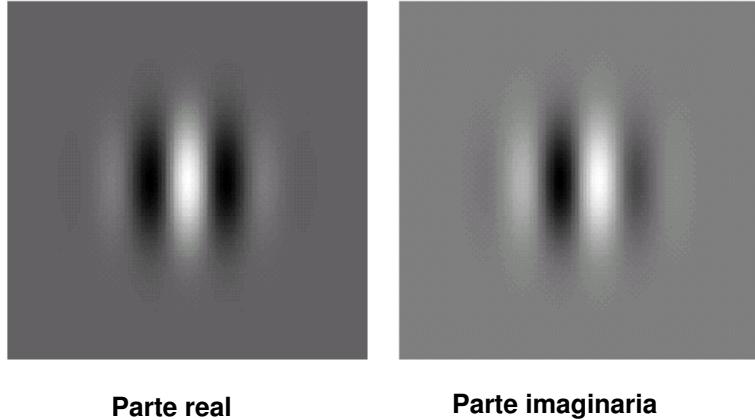


Figura 3.6: Componentes real e imaginário de $\psi(x, y)$.

As funções de Gabor formam um completo, mas não ortogonal, conjunto de bases. Exemplo de uma classe particular de *wavelets* de Gabor 2D são apresentadas na figura (3.7). Usando estas bases, expande-se o sinal que fornece uma descrição de freqüência localizada. Isto faz com que as funções Gabor sejam bem reconhecidas pela ótima localização de tempo-freqüência (Mallat, 1999).

Uma classe de funções auto-similares, nomeadas como *Gabor wavelets* são geradas a partir da função $\psi(x, y)$, a *wavelet* de Gabor mãe. Então o dicionário de filtros auto-similares pode ser obtido mediante dilatações e rotações apropriadas de $\psi(x, y)$ através da função *generatrix*:

$$\psi_{mn}(x, y) = a^{-m}\psi(x', y'), a > 1, m = 1 : M, n = 1 : N. \quad (3.31)$$

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = a^{-m} \begin{bmatrix} \cos\theta_n & \sin\theta_n \\ -\sin\theta_n & \cos\theta_n \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}, \theta_n = n\pi/N. \quad (3.32)$$

onde a^{-m} é o parâmetro de escala que tem como objetivo assegurar que a energia é independente de m .

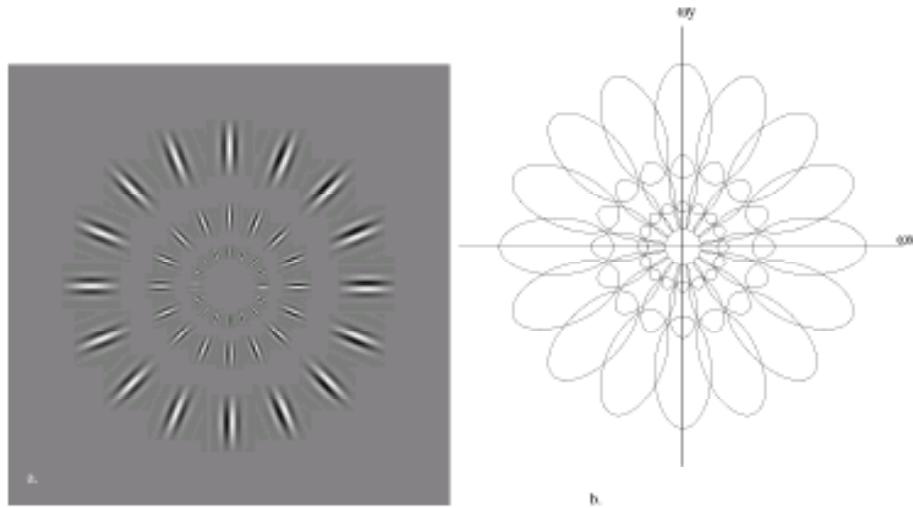


Figura 3.7: (a) Conjunto de *wavelets* Gabor (comprimento de banda octal de 1.5) (b) sua cobertura no plano de freqüência espacial (Lee, 1996).

As *wavelets* de Gabor podem ser interpretadas como um conjunto de funções Gabor com distintos centros de freqüência e orientações (figura 3.7). O tamanho ou o comprimento de banda das *wavelets* de Gabor é também controlado por θ . Pelo fato das *wavelets* Gabor serem simétricas, só precisa-se especificar o valor de θ para montar um espaço uniformemente amostrado em $[0, \pi]$. Dessa forma o conceito de localização das *wavelets* de Gabor estendeu-se para o tempo, freqüência e orientação.

A não ortogonalidade das Gabor *wavelets* implica que existe informação redundante nas imagens filtradas. Isto é verificado na figura 3.7 onde as elipses (correspondente a cada filtro) possuem intersecção entre si. Manjunath (Manjunath and Ma, 1996) define uma estratégia para reduzir esta redundância, que segue os seguintes princípios (Rubner and Tomasi, 2000):

- *Separação uniforme na orientação.* Assumindo a simetria rotacional, todos os filtros em uma escala específica deveriam ter o mesmo desvio-padrão angular (σ_v) e deveriam estar igualmente espaçados no eixo de orientação.
- *Separação exponencial na escala.* Os comprimentos dos filtros deveriam aumentar exponencialmente com distância a partir do centro do plano (u, v) . Isto quer dizer que a diferença entre uma escala e outra está dada pela multiplicação por um parâmetro escalar.
- *Cobertura continua do espaço de freqüência.* O contorno de dois filtros vizinhos tocam um a outro nos eixos de escala e orientação.

Essa estratégia para reduzir redundância está definida por U_l e U_h que denotam os centros de freqüência de interesse superior e inferior. Sejam K o número de orientações e S o número de escalas na decomposição multiresolução, então a estratégia consiste em assegurar que o

suporte da magnitude do pico médio (*half-peak*) dos filtros de resposta no espectro de freqüência toquem-se uns aos outros, como é mostrado na figura 3.8. O processo anteriormente descrito é produto das seguintes fórmulas para o cálculo dos parâmetros dos filtros σ_u e σ_v (e assim σ_x e σ_y):

$$a = (U_h/U_l)^{\frac{1}{S-1}}, \quad (3.33)$$

$$\sigma_u = \frac{(a-1)U_h}{(a+1)\sqrt{2\ln 2}}, \quad (3.34)$$

$$\sigma_v = \operatorname{tg} \left(\frac{\pi}{2K} \right) \left[U_h - 2\ln \left(\frac{2\sigma_u^2}{U_h} \right) \right] \left[2\ln 2 - \frac{(2\ln 2)^2 \sigma_u^2}{U_h^2} \right]^{-\frac{1}{2}}, \quad (3.35)$$

onde $W = U_h$ e $m = 0, 1, \dots, S-1$. Para o presente trabalho foram utilizadas $U_h = 0.3$ e $U_l = 0.04$, o limite superior para a freqüência espacial é de 0.5. A figura 3.8 mostra a localização e comprimento dos filtros resultantes de Gabor, usando quatro escalas e seis orientações.

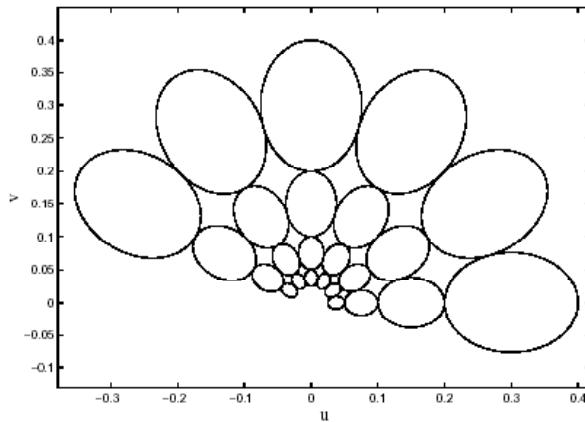


Figura 3.8: Filtros de resposta ortogonais, após aplicar a técnica de eliminação de redundância (Manjunath and Ma, 1996).

3.6 Considerações Finais

Neste capítulo foi mostrada a transformada *wavelet* considerando sua evolução a partir da transformada de Fourier e a sua posterior formalização matemática.

É importante observar que a transformada *wavelet* permite a análise local do sinal. Isso ocorre devido às funções base de suporte compacto. Especial atenção tem-se dado ao uso da função Gabor como *wavelet* mãe, que finalmente é a que melhor desempenho teve nos resultados deste trabalho, justificando-se pelo fato dos filtros de Gabor constituírem bons detectores

de características em imagens (Manjunath and Ma, 1996), (Jain and Healey, 1998), além de possibilitarem minimizar a incerteza das articulações bi-dimensionais no espaço e freqüência (princípio de incerteza de Heisenberg) (Daugman, 1988). Outro aspecto que motiva os pesquisadores a utilizar os filtros de Gabor em análise de imagens é o comportamento biológico destes, pois pesquisas em psicologia mostram que reações de simples células no córtex visual podem ser modeladas mediante as funções Gabor (Daugman, 1988) (Petkov and Kruizinga, 1997).

O uso das *wavelets* neste trabalho demonstra o grande potencial desta técnica na extração de características para sistemas CBIR, no caso das imagens médicas, pelo fato de apresentarem regiões com texturas, as *wavelets* de Gabor são as que melhor têm se comportado, como será apresentado nos capítulos seguintes.

Wavelets e Extração de Características

4.1 Considerações Iniciais

Como foi discutido no capítulo 3, a análise de multiresolução de *wavelets* apresenta muitas funcionalidades vantajosas que podem ser aproveitadas no processo de extração de características das imagens. Além da definição da função *wavelet* adequada ao processamento das imagens médicas, também é importante estabelecer as técnicas de seleção e extração de características a serem utilizadas, como foi tratado no capítulo 2 desta dissertação.

Neste capítulo são apresentados aspectos a serem considerados na geração do conjunto de vetores de características para o conjunto de imagens médicas utilizando os filtros de *wavelets*. Isto é, são considerados os processos de transformação, extração de características e avaliação do conjunto de vetores gerados. A seguir são descritas algumas atividades e considerações referentes a esses passos no processo de extração de características considerando abordagens de multiresolução de *wavelets*.

4.2 Similaridade Semântica das Imagens

Na maioria dos sistemas de recuperação de imagens por conteúdo, o usuário não está interessado nas características de baixo nível das imagens (cor, textura e forma), isso porque o usuário

geralmente pensa em termos de conceitos semânticos. Por exemplo, alguma imagem que ele se lembra de já ter analisado e que deseja recuperar do banco de imagens, ou mesmo buscar as imagens mais parecidas com uma determinada imagem sob análise.

Nesse caso é preciso desenvolver um sistema que tenha a capacidade de armazenar e recuperar as imagens em termos de suas propriedades semânticas. Em (Bradshaw, 2000), apresenta-se uma abordagem da recuperação de imagens baseada na semântica aplicando um enfoque probabilístico. Já o trabalho de (Vailaya, 2000) aproveita as características de baixo nível para aplicá-los no desenvolvimento da categorização semântica das imagens.

Wang (Wang, 2000) desenvolveu um enfoque sensível à semântica visando solucionar problemas de busca em uma base de imagens de propósito geral. Os métodos de classificação semântica são usados para categorizar imagens, tal que esses métodos de busca semântica sejam aplicáveis a cada uma das categorias. Além disso, o sistema pode limitar o alcance da busca a um subconjunto da base original para facilitar um processamento rápido. No caso de imagens médicas o método poderia classificar em “raios-X”, “MRI”, “patologia”, “grafos”, “micro-arranjo”, etc. Dessa forma, o grande problema é procurar um método adequado de extração de características junto a uma métrica de casamento (*matching*) que tenha a capacidade de diferenciar cada uma das classes semânticas.

Uma abordagem recente que tenta fazer uma classificação semântica das imagens é apresentada em (Wang et al., 2001). Nesse trabalho o processo de segmentação aplica a transformada *wavelet* a pequenos blocos de 4×4 para extrair um vetor de características para cada bloco. Em seguida utiliza o algoritmo de clusterização *k-means* (Theodoridis, 1999) diminuindo significativamente o cálculo computacional da segmentação.

No presente trabalho, o problema de *definir o domínio semântico de recuperação de imagens* é dado pela semelhança das imagens médicas em conjuntos de exames de mesmo tipo. Por exemplo cortes de cabeça, de bacia, entre outros (frontal, vertical, transversal), podem formar um grupo semântico, e cada vez que seja feita uma consulta o sistema deve fornecer como resposta as imagens correspondentes ao mesmo grupo semântico.

4.3 Propriedades e Características das Imagens Médicas

Como o cérebro faz para transformar as imagens da retina para uma representação mais eficiente e útil que torna explícitos os objetos, formas, movimentos, etc. que estão presentes no ambiente? Dados neurofisiológicos indicam que os aspectos mais complexos das formas dos objetos são extraídos em forma hierárquica através das áreas do cortex visual. Recentemente, Felleman e sua equipe (Xiao et al., 2003) da Universidade de Texas-Houston, descobriram como o cérebro percebe as cores através de células nomeadas como “tiras finas” (*thin stripes*) localizadas na região V2 do cérebro. Tais células contém mapas funcionais onde os estímulos de cor

são representadas mediante um pico de ativação de resposta. Uma análise do funcionamento do cérebro com uma abordagem estatística é apresentada em (Rao et al., 2002). O maior interesse no presente trabalho está em conhecer a estrutura das imagens médicas, que constituem minúsculas frações do espaço de todas as possíveis imagens.

As imagens naturais (no nosso caso, médicas) contêm estruturas de passa banda orientadas e localizadas, que não podem ser caracterizadas em termos de correlações lineares ou duais. As estruturas localizadas em imagens naturais são caracterizadas, em termos de Fourier, mediante seu espectro de fase. Esta distribuição localizada e compacta da energia nas imagens, implica que elas tenham uma “estrutura esparsa” (Olshausen, 2002), isto quer dizer, que elas contêm componentes de alta freqüência que diminuem a inata redundância (estrutura fractal), o qual permitirá que uma imagem qualquer possa ser representada com um relativamente pequeno número de descritores.



Figura 4.1: Imagem com áreas de textura que identificam regiões distintas

No caso das imagens médicas, elas cumprem com as mesmas condições descritas nos parágrafos anteriores, estas características das imagens naturais são muito bem aproveitadas pelo processo de transformação baseado em *wavelets*, conseguindo reduzir a informação de forma ótima (dada a estrutura esparsa das imagens), e a estrutura fractal dos tecidos permite identificar distintas regiões, como se observa na figura 4.1. Isto também é aproveitado pelos profissionais médicos para identificar anomalias dentro de um órgão ou região da imagem.

4.4 Escolha da Wavelet Analisadora

Como foi explicado no capítulo 3, a análise de multiresolução de *wavelets* (Mallat, 1989) é uma estratégia que se mostra bastante adequada para a extração das características das imagens pelas seguintes razões:

- poucos coeficientes representam bem a imagem,
- armazena informações de arestas e bordas,
- praticamente não dependem da resolução original,
- a transformada tem custo linear, sendo dependente do tamanho da imagem.

Em geral, o foco de muitas pesquisas recentes em *wavelets* é criar uma função de *wavelet* mãe para obter uma descrição útil, eficiente e informativa do sinal de interesse. Não é fácil planejar um procedimento uniforme para desenvolver a melhor *wavelet* mãe ou transformada *wavelet* para um determinado sinal. Mas, considerando as características gerais das funções *wavelets*, é possível determinar quais das funções *wavelets* são mais adequadas para uma determinada aplicação.

Características de representação no tempo e escala, assim como o fato de serem compactos e de possuírem energia finita fazem com que as funções *wavelets* sejam aplicáveis para analisar não só sinais estacionários, mas também para decompor sinais dinâmicos, de tempo variável, ou sinais transientes.

Do ponto de vista prático, existem dois tipos de decomposição baseada em *wavelets*: a primeira que é denominada como redundante (*Continuos Wavelet Transform - CWT*) ou *wavelets* de *frames*, e a segunda é denominada como não redundante (bases de *wavelets* ortogonal, semi-ortogonal ou bi-ortogonal). As *wavelets* redundantes são preferencialmente usadas em análise de sinais, extração e detecção de características. Já as *wavelets* não redundantes são mais adequadas quando precisa-se realizar algum tipo de redução de informação ou quando a ortogonalidade da representação é um fator de importância. Porém, escolher entre estas duas opções não é simples devido às considerações computacionais que devem ser feitas (Unser and Aldroubi, 1996).

As propriedades das *wavelets* podem ser muito bem aproveitadas em várias aplicações, Unser (Unser and Aldroubi, 1996) define algumas destas propriedades e o contexto de aplicação em que elas podem ser utilizadas: (a) *wavelets* como bancos de filtros, (b) *wavelets* como filtros multiescala de casamento, (c) *wavelets* e localização Tempo-Freqüência, (d) bases de *wavelets* e (e) *wavelets* como modelo de percepção.

Existe uma variedade de funções *wavelets* que têm comportamento distinto de acordo com o domínio de imagens. As *wavelets* Daubechies (Daubechies, 1992) apresentam-se como as que melhor caracterizam os detalhes da imagem através das *wavelets* mãe que ficam nas sub-bandas de baixa resolução (os coeficientes de *wavelets*) enquanto que as *wavelets* pai caracterizam a tendência do sinal (Wang, 2000). Estas propriedades podem ser aproveitadas para caracterizar uma imagem e gerar o espaço de características aplicando descritores estatísticos sobre esses novos espaços.

Como foi apresentado na seção 4.3 as imagens médicas apresentam um conjunto de regiões texturadas não homogêneas que precisam ser caracterizadas. Na seção 3.5 definiram-se os filtros de Gabor que estão considerados como os que melhor caracterizam a textura pelo fato que células simples do cortex visual podem ser modelados mediante esses filtros, sendo uma das tarefas principais nessa abordagem a determinação dos parâmetros apropriados. O uso do conjunto de filtros com parâmetros ótimos não só acrescenta a eficiência computacional, mas também, extrai maior informação significativa. Na seção 3.5.1 apresentou-se a determinação dos parâmetros dos filtros de Gabor baseados no trabalho de Manjunath (Manjunath and Ma, 1996) o qual é utilizado nesse trabalho para gerar o espaço de características.

4.5 Geração do Espaço de Características

A parte crítica do processo de recuperação de imagens é definir o espaço de características a ser utilizado para realizar o casamento entre atributos similares. Espaços de características comuns incluem bordas, contornos, superfícies e outras características salientes tais como esquinas, intersecção de linhas e pontos de alta curvatura, da imagem.

O espaço de características é a representação dos dados que serão usados no processo de recuperação de imagens por conteúdo. A eleição do espaço de características determina o que será casado no momento da busca de imagens mediante a métrica de semelhança que determina quão parecidas são duas imagens.

No caso da aplicação da transformada *wavelet* para a geração do espaço de características, tem-se que efetuar uma análise estatística em cada uma das sub-bandas, onde cada elemento do vetor poderia corresponder a uma descrição de cada subimagem, de maneira que o conteúdo presente na imagem seja corretamente descrito pelo vetor. Para efetuar essa análise estatística pode-se utilizar os coeficientes estatísticos descritos em 2.4.1, como a energia, desvio-padrão, etc.

Uma descrição mais completa, mas computacionalmente cara, é feita através de uma matriz de co-ocorrência (no caso de características de textura). Uma vez que a matriz de co-ocorrência é calculada são determinados alguns parâmetros que permitem estabelecer o relacionamento da textura em uma imagem como foi especificado na seção 2.4.2.1.

No caso de trabalhar com momentos invariantes, uma forma de gerar o vetor de características é calcular n níveis de decomposição da *wavelet* em escalas $2^j, 1 \leq j \leq n$, para então aplicar os momentos invariantes $\phi_i^j, 1 \leq i \leq 7$ (momentos de Hu), a cada uma das escalas. Para o caso de gerar 4 níveis de *wavelets*, então ter-se-iam 28 números reais como assinatura para cada imagem (Do et al., 1999).

4.6 Medidas de Distância

Na procura de imagens visualmente similares, dada uma imagem de consulta, em um sistema de recuperação de imagens é importante levar em consideração dois aspectos: uma representação apropriada das características visuais das imagens e uma medida que determine o grau de similaridade ou dissimilaridade das imagens de resposta. Considerando que não existe nenhuma anotação textual da imagem, as características que podem ser utilizadas são descrições do conteúdo da imagem que podem ser representados pela cor, textura e forma. Estas características mudam sobre uma mesma imagem, por variações inerentes da aparência da superfície e como resultado de mudanças na iluminação, sombreamento, entre outros. Dessa forma, a aparência de uma região é melhor descrita pela *distribuição* das características, mais do que pelos vetores de características individuais.

Medidas de similaridade, baseadas em estimativas empíricas da distribuição das características foram desenvolvidas e utilizadas para diferentes tarefas da área de visão computacional, por exemplo em classificação, recuperação de imagens e segmentação. O livro de (Rubner and Tomasi, 2000) pode ser considerado como uma boa referência sobre métricas, onde são classificadas e analisadas algumas das medidas de similaridade mais comuns.

O primeiro passo importante é definir uma medida de distância entre dois objetos. Essa medida de distância está denotada por $d()$ e é definida como:

Sejam dois objetos, S e Q pertencentes a um mesmo domínio, a distância (ou dissimilaridade) entre os dois objetos é denotada por

$$d(S, Q) \quad (4.1)$$

Por exemplo, se os objetos tem sido caracterizados por vetores de características de igual tamanho, a distância $d()$ poderá ser a distância Euclidiana entre eles (soma da diferença de quadrados):

$$d(S, Q) \equiv \left(\sum_{i=1} (S[i] - Q[i])^2 \right)^{1/2} \quad (4.2)$$

Onde S e Q são dois objetos, $S[i]$ e $Q[i]$ correspondem ao valor da i -ésima dimensão de S e Q respectivamente.

Uma outra medida de distância muito utilizada é conhecida como distância de Mahalanobis (equação B.7), que é usada no processamento de imagens para identificar os valores variantes durante a classificação da imagem.

Albuz (Albuz et al., 2001) propõe uma medida de distância para regiões IRM (*Integrated Region Matching*) que torna mais robusta a métrica quando se tem uma segmentação fraca. Albuz obtém bons resultados gerando características de textura para imagens com textura definida.

4.7 Análise e Avaliação Visual

A análise das características extraídas implica na definição do melhor subconjunto (o conjunto de características candidatas) que sejam as mais representativas da imagem em processo. Por isso é necessário estabelecer uma maneira de avaliar as características que estão sendo extraídas.

Dependendo dos tipos de imagens e das aplicações envolvidas, há uma plethora de características a serem extraídas, de maneira a serem utilizadas em cada situação específica. Por exemplo, um patologista que sabe que núcleos de células de certos tamanhos são importantes para sua análise gostaria que o sistema examinasse uma série de imagens de células e que mostrasse só aquelas que contém o núcleo desejado. Já um radiologista comparando várias imagens de ultra-som gostaria de saber se a textura em uma determinada região anatômica das diversas imagens são similares. Em ambos os casos, o usuário sabe o que esperar como resposta, mesmo que intuitivamente.

Infelizmente, em vários tipos de aplicações, a escolha das características a serem utilizadas não é clara. Por exemplo, “existe alguma patologia difusa nessa imagem de ultra-som?”, “será que esse mamograma tem algum tumor presente?”, “será benigno ou maligno?”. Nestes casos é preciso classificar uma imagem ou subimagem em uma das classes do conjunto. Representar uma imagem mediante um conjunto de valores característicos é equivalente a associá-la com um ponto no espaço de dimensionalidade igual ao número de características no vetor, onde aqueles que têm valores próximos entre si possuirão maior grau de similaridade (Traina et al., 2000), (Ciaccia et al., 1997).

A análise visual do conjunto de características extraídas das imagens por esse trabalho foi realizada utilizando uma ferramenta previamente desenvolvida no GBDI. Esta ferramenta, denominada de *FastMapDB* (Barioni et al., 2002), foi acrescentada com diversas técnicas de visualização como parte do trabalho de mestrado de outro membro do grupo de pesquisa (Rodrigues, 2003). Uma das características importantes do uso desta ferramenta é a visualização direta e intuitiva dos vetores de características multidimensionais mapeados em um espaço 3D. Isto propicia perceber a distribuição destes vetores no espaço bem como a formação dos *clusters* de acordo com o domínio semântico das imagens.

A limitante para o uso do *FastMapDB* encontra-se na geração do espaço de características estar vinculado a uma função de similaridade, a qual deve ser programada no sistema para visualizar o resultado da distribuição dos vetores de características. Como complemento a esta ferramenta, gerou-se grafos do conhecido método de precisão/revocação (*precision/recall*)

(Baeza-Yates and Ribeiro-Neto, 1999) onde a maior área abaixo da curva indica que o extrator discrimina bem os dados, mas gerar esses grafos requer uma grande quantidade de tempo além de se precisar de um especialista para determinar se o conjunto de resposta corresponde ao domínio de interesse.

4.8 Considerações Finais

Neste capítulo foram apresentados alguns aspectos importantes a ser considerados no processo de extração de características mediante a aplicação da transformada *wavelet*; assim como mostraram-se pesquisas desenvolvidas nesta área que visam solucionar os problemas existentes na recuperação de imagens por conteúdo.

Além da geração do vetor de características, a definição de uma métrica será de muita importância para responder às consultas por similaridade, além da indexação do espaço de características em uma estrutura métrica (no presente trabalho será utilizada a *Slim-tree* (Traina et al., 2000)).

Experimentos e Resultados

5.1 Introdução

Neste capítulo, apresenta-se uma descrição detalhada dos experimentos desenvolvidos com as técnicas de extração de características baseada em *wavelets* detalhadas nos capítulos anteriores. Especificamente, descrevem-se os conjuntos de dados usados nos experimentos, as abordagens adotadas na extração de características, e as respectivas métricas de distância utilizadas no casamento de imagens similares, comparações entre as distintas abordagens, a robustez da técnica a algumas mudanças, e a performance na indexação e busca.

5.2 Visão Geral

Conforme mencionamos no capítulo inicial, o principal objetivo deste trabalho é o estudo de técnicas para extrair características baseadas na transformada *wavelet*, com a finalidade de possibilitar uma caracterização das imagens médicas referentes a exames radiológicos e tomográficos. Dessa forma, o sistema cbPACS ganhará uma nova funcionalidade para satisfazer as consultas por similaridade, que desejamos que atenda às necessidades de buscas semânticas sobre as imagens.

Basicamente, foram estudadas duas abordagens para extração de características baseadas na transformada *wavelet*. A primeira, e a mais simples, consiste na aplicação de filtros de Daubechies-2 (Haar) e Daubechies-4 (seções 3.4.2 e 3.4.3), que são filtros simples de imple-

mentar através da técnica de multiresolução. A segunda abordagem consiste na aplicação da função Gabor como *wavelet* mãe (seção 3.5), o que é denominado como transformada *wavelet* de Gabor. Em comparação à primeira abordagem, as *wavelets* de Gabor são consideradas como as melhores de localização no tempo e na freqüência, por conter maior número de parâmetros e, as mais atrativas, por simular reações simples de células no cortex visual.

Após aplicar a transformada *wavelet*, mediante uma abordagem de multiresolução, geram-se novos espaços os quais são aproveitados para realizar uma análise estatística para a extração de características. Medidas estatísticas podem ser utilizadas para extrair características e avaliar a luminosidade da imagem (através da média), a suavidade da imagem (através da entropia), e a uniformidade (através da energia), entre outras. Algumas outras medidas são especificadas na seção (2.4.1).

Finalmente, uma das tarefas mais onerosas é a que se refere ao processo de avaliação. No método tradicional precisa-se classificar o conjunto de imagens indicando as que são relevantes para cada tipo de consulta, e aplicar várias consultas por similaridade sobre o banco de imagens, em seguida um especialista precisa validar as respostas fornecidas pela técnica de recuperação, e assim gerar um "grafo de precisão e revocação" (*precision and recall graph*). Neste trabalho aproveita-se o método desenvolvido por (Rodrigues, 2003), dentro do nosso Grupo de Bases de Dados e Imagens (GBDI-ICMC), que é uma abordagem visual que auxilia na avaliação de espaços de características.

Na figura (5.1) ilustra-se o processo geral de extração de características, onde, dependendo da abordagem, em cada um dos processos internos realizam-se tarefas de acordo com os parâmetros de cada um deles. Assim mesmo, o número de valores gerados para cada vetor de características vai depender de cada abordagem.

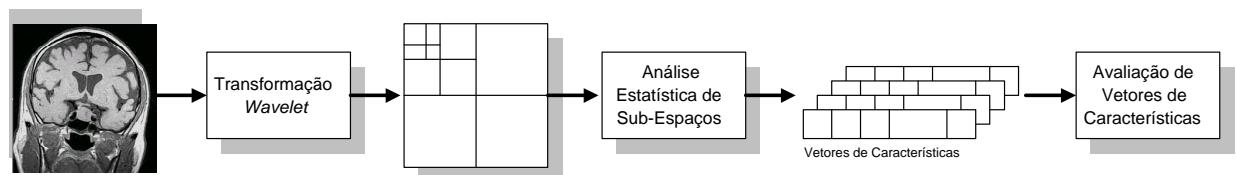


Figura 5.1: Processo geral de extração de características

5.3 Conjunto de Dados

As imagens utilizadas no trabalho correspondem a exames de ressonância magnética (RM) e tomografia computadorizada (CT). Pode-se notar a diferença entre elas pelo fato de terem níveis de contraste distintos na superfície das imagens e mudança nos tons de cinza (figura 5.2). As imagens de ressonância magnética ressaltam melhor os tecidos moles (texturas), enquanto as tomografias computadorizadas refletem melhor os tecidos duros (a parte óssea da imagem).

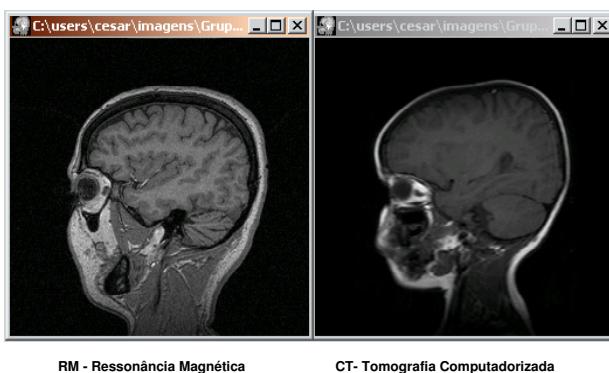


Figura 5.2: Imagens de RM e CT (percebe-se a diferença de tons de cinza)

Os exames correspondem a cortes sucessivos. Isto faz com que as imagens sejam semelhantes a suas consecutivas, mas dependendo da espessura do corte a textura vai mudando de uma imagem para outra, sendo que a similaridade diminui com a separação entre as fatias (figura 5.3).

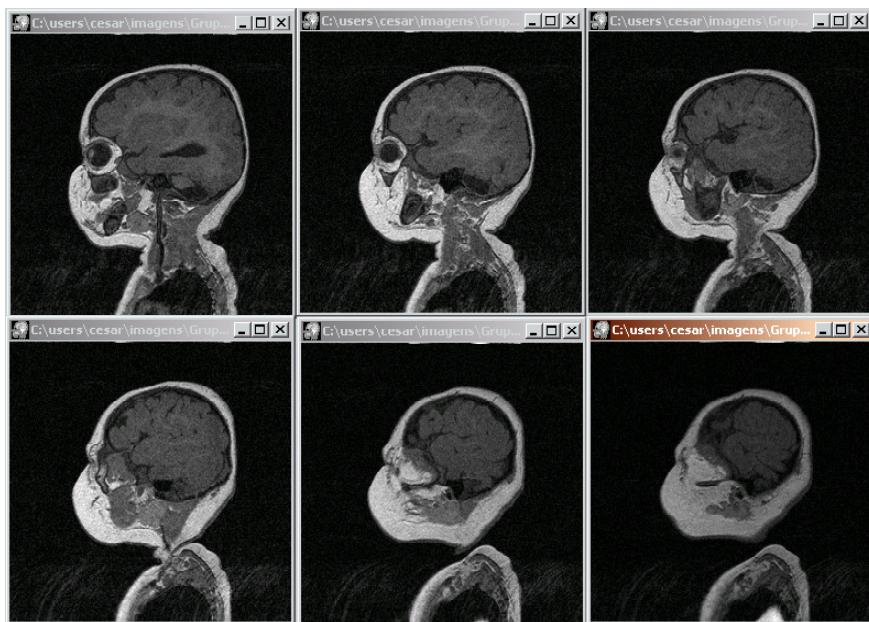


Figura 5.3: Seqüência de imagens RM com espessura distinta nos cortes

As imagens consideradas no domínio de consulta representam partes distintas do corpo humano, dependendo dos exames feitos em um paciente. Por exemplo cortes sagitais e coronais de cabeça, medula, bacia, abdomen, pélvis, tórax, cérebro, espinha, entre outros; alguns exemplos são apresentados na figura 5.4.

O número de imagens por cada exame de paciente é variável. Essas imagens estão armazenadas em formato DICOM (que difere entre as modalidades de exames), onde as dimensões da maioria delas é de 256×256 , mas também têm-se imagens de 512×512 . Cada uma das abordagens seguidas toma um subconjunto aleatório destas imagens para realizar testes.

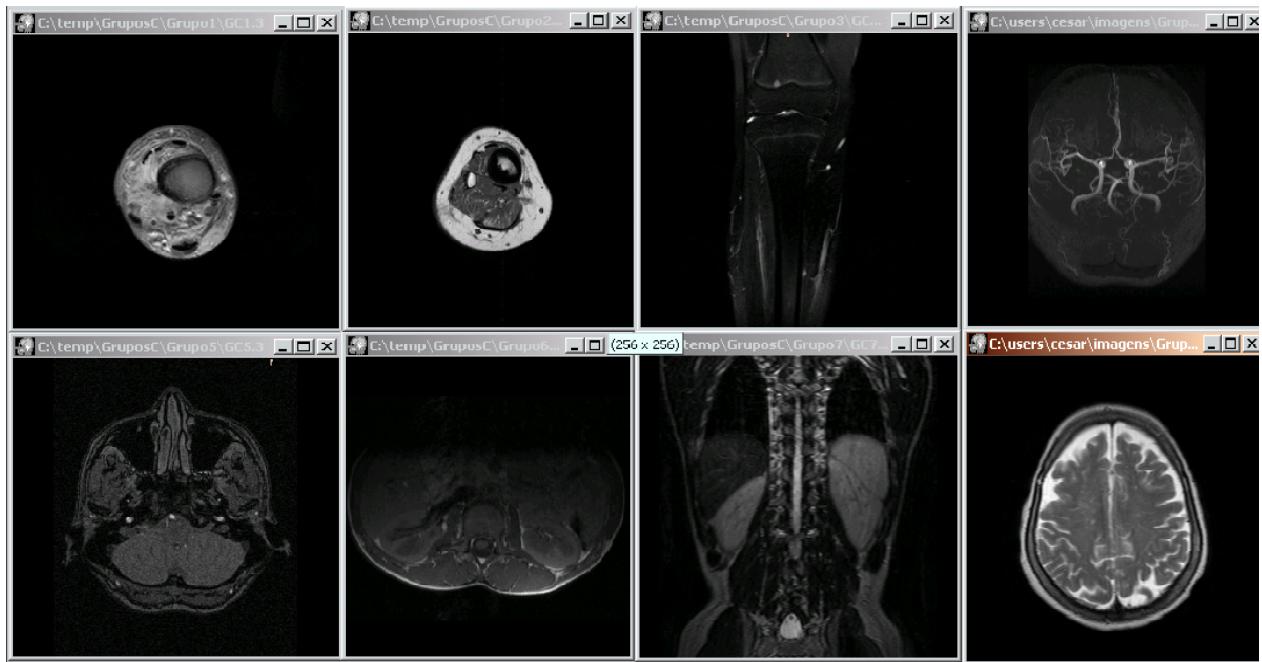


Figura 5.4: Exemplos de imagens da base utilizada

5.4 Grupos de Teste

Para realizar os testes das diferentes técnicas desenvolvidas neste trabalho e avaliar a performance de cada uma delas definiu-se três grupos de teste: *GrupoBásico* que é utilizado para analisar o comportamento do caracterizador, *GrupoIntermediário* que define um maior número de imagens dentro das quais encontram-se algumas denominadas como imagens ruido, e *GrupoMaior* o qual foi montado com o objetivo de simular um grande banco de imagens e analisar o comportamento real das consultas por similaridade.

5.4.1 Grupo Básico

Este subconjunto de imagens, denominado como *GrupoBásico*, tem como objetivo realizar uma simples avaliação do comportamento do extrator de características. Para isso foram selecionadas cinco grupos de imagens, as que contém imagens previamente escolhidas que possuem similaridade entre elas. Na figura 5.5 mostra-se uma imagem de cada grupo e na tabela 5.1 detalha-se as características do grupo básico.

5.4.2 Grupo Intermediário

O grupo intermediário (*GrupoIntermediário*) é constituído por um grupo representativo de 500 imagens médicas, divididos em 8 subgrupos. Os sete primeiros subgrupos, formados por 30



Figura 5.5: Imagens do *GrupoBásico* de teste, cada imagem é uma representativa de cada grupo

Grupo	Descrição	Nro. de Imagens
A	Cortes sagitais de RM - Cabeça	20
B	Cortes coronais de RM - Abdómem	21
C	Cortes axiais de RM - Bacia	21
D	Corte sagital de RM - Cabeça	16
E	Angiograma	06
	Total Imagens	84

Tabela 5.1: Conjunto de imagens de teste - *GrupoBásico*

imagens cada, representam imagens previamente selecionadas considerando sua semelhança. Na figura 5.6 observa-se uma imagem representativa de cada subgrupo.



Figura 5.6: Imagens do *GrupoIntermediário* de teste, cada imagem é uma representativa de cada grupo

Além dos sete subgrupos acima descritos, o oitavo subgrupo é composto de 290 imagens médicas, a qual forma uma coleção de imagens complementares aos sete primeiros subgrupos e que não mantêm características comuns necessariamente. Este último subgrupo é adicionado como subgrupo de ruído para avaliar a robustez dos extractores. Na tabela 5.2 resumem-se as características do *GrupoIntermediário*.

Grupo	Descrição	Nro. de Imagens
A	Corte sagital de RM - Cabeça	30
B	Corte sagital de RM - Abdómem	30
C	Corte axiais de RM - Bacia	30
D	Corte axial de RM - Cabeça	30
E	Angiograma	30
F	Corte sagital de RM - Espinha	30
G	Corte sagital de RM - Cabeça	30
H	Diferentes imagens de domínios distintos	290
	Total Imagens	500

Tabela 5.2: Conjunto de imagens de teste - *GrupoIntermediário*

5.4.3 Grupo Maior

O grupo maior (*GrupoMaior*) tem por objetivo simular um banco de imagens médicas real, considerando um volume de imagens aceitável, que ofereça resultados confiáveis. O número total é de 5631 imagens médicas de diversos cortes do corpo humano, sendo as imagens de tamanho 256×256 e de 512×512 . Um resumo do grupo é mostrado na tabela 5.3.

CD	Nro. Exames	Imagens 512×512	Imagens 256×256
MV300CD260	14	244	2084
MV300CD262	19	274	3029
Total	31	518	5113
		Total Imagens	5631

Tabela 5.3: Conjunto de imagens de teste - *GrupoMaior*

5.5 Características Baseadas na Distribuição da Cor

5.5.1 Motivação

As propriedades das *wavelets* permitem desenvolver uma série de técnicas que visem caracterizar as imagens, entre estas propriedades temos (Albuz et al., 2001):

- *Localização*: Cada componente *wavelet* é localizada simultaneamente no tempo e na freqüência. Pelo que as *wavelets* podem casar uma ampla extensão de componentes do sinal desde transitórios até harmônicos.
- *Multiresolução*: Os componentes de *wavelets* esticam-se e dilatam-se para a análise dentro de conjuntos aninhados de escalas. Isto permite que a transformada case tanto com as estruturas do sinal de curta como de longa duração.
- *Compressão*: As transformadas *wavelet* de sinais e imagens do mundo real, tendem a ser esparsas. Como resultado a distribuição dos coeficientes de *wavelets* de sinais e imagens do mundo real podem ser comprimidas.

No presente trabalho realizamos testes visando aproveitar estas características aplicando descritores estatísticos sobre a distribuição das cores (nesse caso considerou-se tons de cinza pelas imagens que apresentam estas características, mas o método pode ser estendido para imagens coloridas).

5.5.2 Algoritmo de Decomposição

Existem duas formas comuns em que as *wavelets* podem ser usadas para transformar os valores dos pixels dentro de uma imagem. Cada uma destas transformações é uma generalização da transformada *wavelet* uni-dimensional descrita na seção 3.4.4 (Stollnitz et al., 1996).

A primeira transformada é denominada de *decomposição padrão* a qual consiste em aplicar a transformada uni-dimensional de *wavelet* em cada linha da matriz de pixels. Em seguida, aplica-se a transformada em cada coluna da matriz. No caso de aplicar os filtros de Haar (equações 3.23 e 3.24), esta operação fornecerá a média dos valores com coeficientes de detalhe para cada linha/coluna. O segundo tipo é denominado de *decomposição não padrão* que alterna as operações entre linhas e colunas. Dessa duas abordagens, a padrão é considerada a mais fácil de implementar enquanto que a não padrão se apresenta como levemente mais eficiente de calcular. Para uma imagem de $m \times m$, a decomposição padrão requer de $4(m^2 - m)$ operações de atribuição, enquanto a decomposição não padrão requer só $\frac{8}{3}(m^2 - 1)$ operações de atribuição (Stollnitz et al., 1996).

No caso de aplicar os filtros de *wavelets* de Daubechies-4 (equações 3.25 e 3.26) o processo é similar ao descrito nos parágrafos anteriores, que consideram os filtros de Haar. Mas repare que os filtros de Daubechies-4 possuem quatro componentes, neste caso para fazer a transformação deve-se considerar cada quatro valores (pixels) para efetivar operações de filtragem. Quando chegar no final ficarão os dois últimos valores, faltando dois valores para completar o quarteto, geralmente considera-se nesta última operação completar o quarteto com os dois primeiros valores. Após aplicar os filtros de *wavelets* geram-se novos valores entre os quais têm-se valores negativos. Para efeitos de visualização, é preciso mapear estes valores para o domínio de (0 a 255).

5.5.3 Geração dos Vetores de Características

Uma cuidadosa observação dos subespaços de *wavelets* (figura 5.7) indica que existe uma relação hierárquica entre os distintos subespaços, o que forma uma estrutura piramidal de subimagens com várias resoluções correspondentes a diferentes escalas. Por exemplo, o subespaço LLLH tem informação similar do que o subespaço LH. Isto porque ambos subespaços são gerados pelo mesmo filtro, mas em escalas distintas.

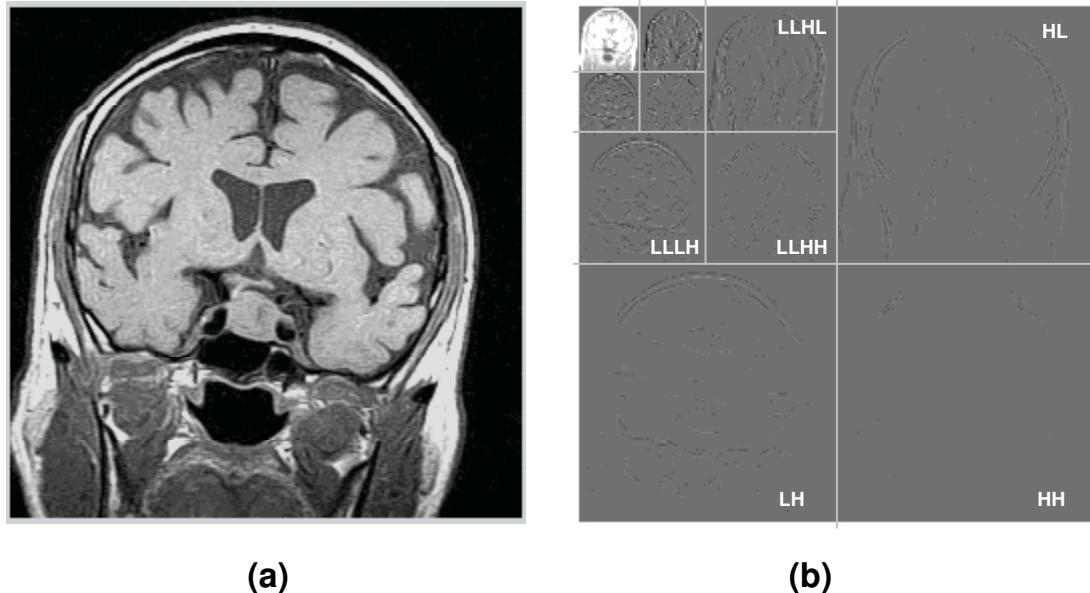


Figura 5.7: (a) Imagem original, (b) Decomposição da imagem após três níveis de transformação com filtros Daubechies-4, o contraste e brilho da imagem foram mudados com a finalidade de ressaltar os detalhes.

O cálculo dos vetores de características está baseado nas propriedades de energia dos subespaços de *wavelets*. Para cada imagem, depois de aplicar transformada *wavelet*, é calculado o valor da energia total do subespaço de *wavelets*, de acordo com a equação 2.4.

Nos testes realizados aplicaram-se três níveis de decomposição, gerando uma subimagem de baixa freqüência, e nove subespaços de *wavelets* (cada três nas direções horizontal, vertical e diagonal) que correspondem a subimagens direcionais de alta freqüência. Aplicando o processo descrito no parágrafo anterior, o vetor de características é composto de nove valores de energia total de cada subespaço de *wavelets* de alta-freqüência, como é mostrado na figura 5.8.

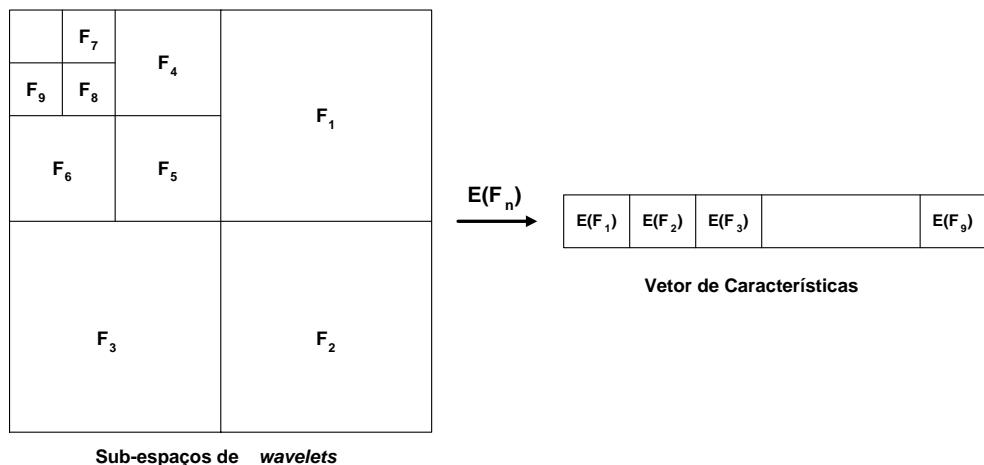


Figura 5.8: Geração do vetor de características de uma imagem baseada nos subespaços de *wavelets* considerando três níveis de decomposição e a função $E(F_n)$ corresponde ao extrator aplicado sobre o subespaço n

5.5.4 Visualização dos Vetores de Características

Para a visualização dos vetores de características é utilizada a ferramenta FastMapDB (Barioni et al., 2002) desenvolvida no GBDI-ICMC-USP, que considera uma redução de dimensionalidade para o espaço euclideano 3D e utiliza como métrica de distância a função *Euclideana*. Na figura 5.9 visualiza-se a distribuição dos grupos no espaço 3D gerados com filtro de *wavelets* Daubechies-4 e considerando o extrator de energia total.

O segundo teste foi desenvolvido considerando o grupo de teste *GrupoIntermediário*, sobre o qual foi aplicada a *wavelets* Daubechies-4 e considerada a distância euclideana como métrica para recuperar os vetores em ordem de similaridade. Foram realizadas uma média de 5 consultas para cada um dos 7 primeiros subgrupos de imagens dentro do conjunto de 500 imagens, totalizando 35 consultas, considerando os três extratores, totalizando um total de 105 consultas realizadas. Na figura 5.10 mostra-se o resultado das consultas em um grafo de precisão e revocação, onde o extrator de entropia forneceu melhores resultados.

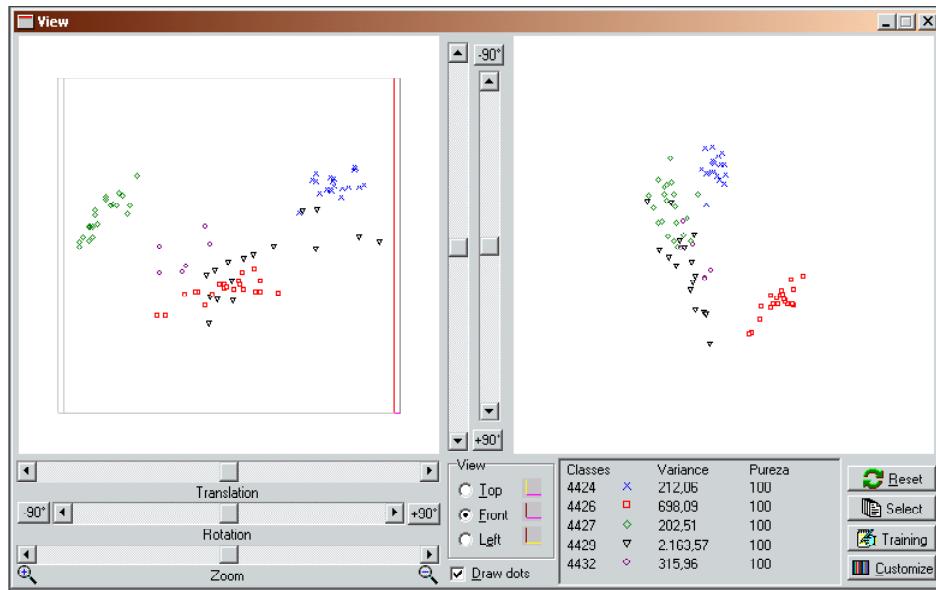


Figura 5.9: Visualização no FastMapDB correspondente a cinco grupos de imagens utilizando como extrator a energia e filtros de *wavelets* Daubechies-4

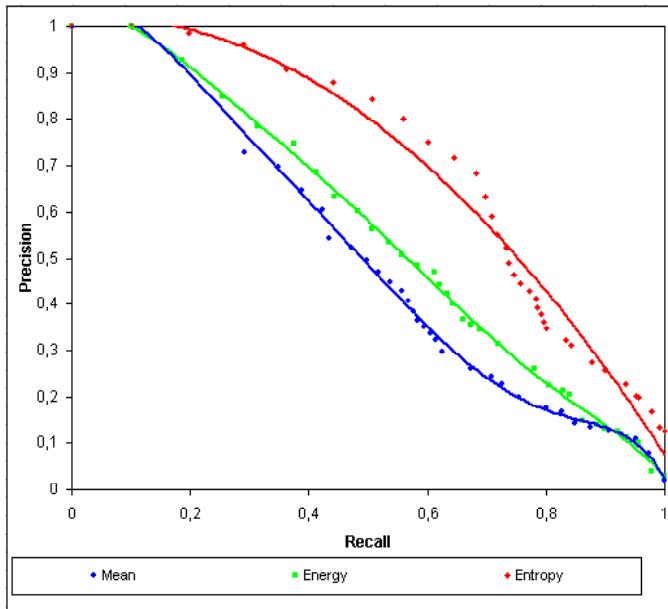


Figura 5.10: Grafo de precisão e revocação das *wavelets* Daubechies-4

5.5.5 Aplicando outros Extratores

Além da energia total, que tem como objetivo avaliar a uniformidade dos pixels da imagem, existem alguns outros extratores que permitem avaliar outras propriedades da imagem, como por exemplo a média do brilho da imagem é dada pela média (equação 2.2), a suavidade da imagem é dada pela entropia (equação 2.5).

No mesmo sentido que procedeu-se com a energia, aplicou-se cada extrator em testes sobre os respectivos subespaços os quais geraram um vetor de características de dimensão nove, con-

siderando três níveis de decomposição. Na figura 5.11 mostra-se o resultado após aplicar (a) média e (b) entropia.

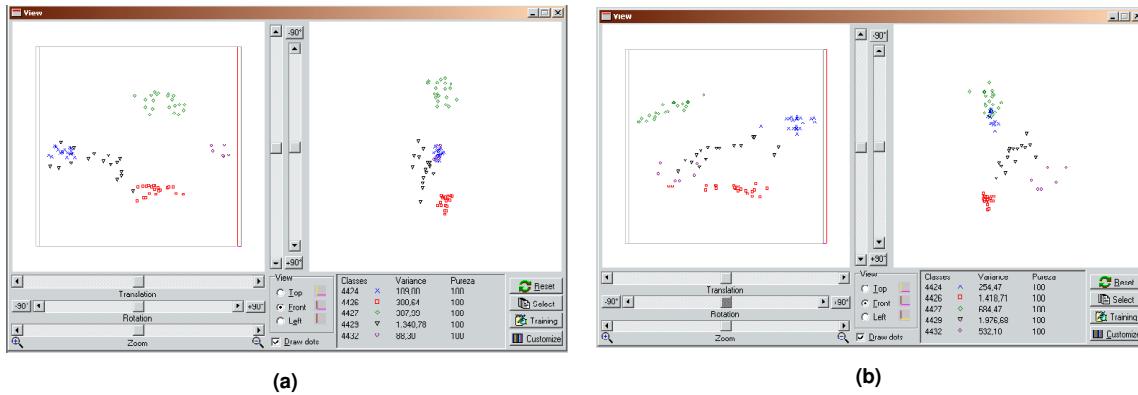


Figura 5.11: Visualização correspondente aos extractores (a) média e (b) entropia, com filtros de *wavelets* Daubechies-4

Esses testes também podem ser feitos utilizando os filtros de *wavelets* de Haar e aplicar sobre os subespaços de *wavelets* os mesmos extractores descritos para Daubechies-4. O resultado é apresentado na figura 5.12, onde (a) corresponde a energia, (b) média e (c) entropia.

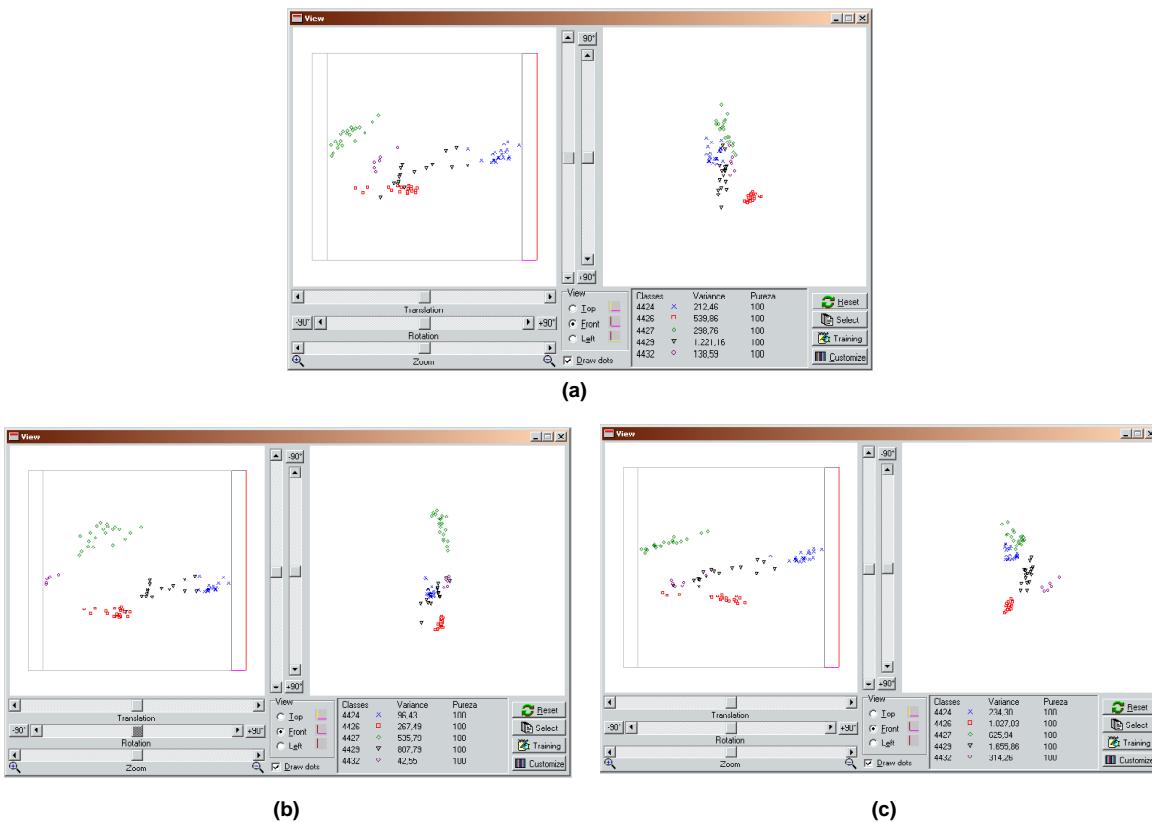


Figura 5.12: Visualização correspondente ao *GrupoBásico* aplicando filtros de *wavelets* Haar e extractores (a) energia, (b) média e (c) entropia.

5.6 Características Baseadas na Textura da Imagem

5.6.1 Motivação

As funções de Gabor 2D são uma rede de senoidais complexas moduladas mediante um sinal Gaussiano 2D no domínio do espaço e translações de funções Gaussiano no domínio de freqüência (Bovik et al., 1990). Daugman (Daugman, 1988) demonstrou que as funções de Gabor são as únicas funções que simultaneamente conseguem o limite inferior do princípio de incerteza. Estas funções podem ter várias formas, comprimentos de banda, centros de freqüência e orientações mediante o ajuste de parâmetros adequados e tem sido muito úteis na análise de imagens com textura que possuem alta freqüência específica ou orientações características (Chitre and Dhawan, 1999).

Neste trabalho, a abordagem considerada na extração de características está baseada no trabalho de Manjunath (Manjunath and Ma, 1996). O embasamento matemático desta técnica foi desenvolvida na seção 3.5.

5.6.2 Extração do Vetor de Características

Aplicando convolução da imagem com os filtros de Gabor, a transformada *wavelet* de Gabor da imagem $I(x, y)$ está definida como (Manjunath and Ma, 1996):

$$W_{mn}(x, y) = \int I(x', y') \psi_{mn}(x - x', y - y') dx' dy', \quad (5.1)$$

Assume-se que as regiões locais de textura são espacialmente homogêneas, e a média μ_{mn} e o desvio-padrão σ_{mn} da magnitude dos coeficientes da transformação são usados para representar cada imagem e assim classifica-las para uma posterior indexação e recuperação:

$$\mu_{mn} = \int \int |W_{mn}(xy)| dxdy, \quad (5.2)$$

$$\sigma_{mn} = \sqrt{\int \int (|W_{mn}(x, y)| - \mu_{mn})^2 dxdy} \quad (5.3)$$

O vetor de características é construído usando μ_{mn} e σ_{mn} como componentes característica. Nos experimentos foram utilizadas duas escalas ($S = 2$) e seis orientações ($K = 6$), resultando um vetor de características com 24 componentes numéricos.

$$\bar{f} = [\mu_{00}, \sigma_{00}, \mu_{01}, \dots, \mu_{15}, \sigma_{15}] \quad (5.4)$$

5.6.3 Desempenho do Extractor

Os vetores extraídos são visualizados utilizando a ferramenta *FastMapDB* e usando como métrica a distância Euclideana. A distribuição dos vetores é apresentada na figura 5.13 onde (a) corresponde ao grupo de teste *GrupoBásico* e (b) ao grupo de teste *Grupo Intermediário*.

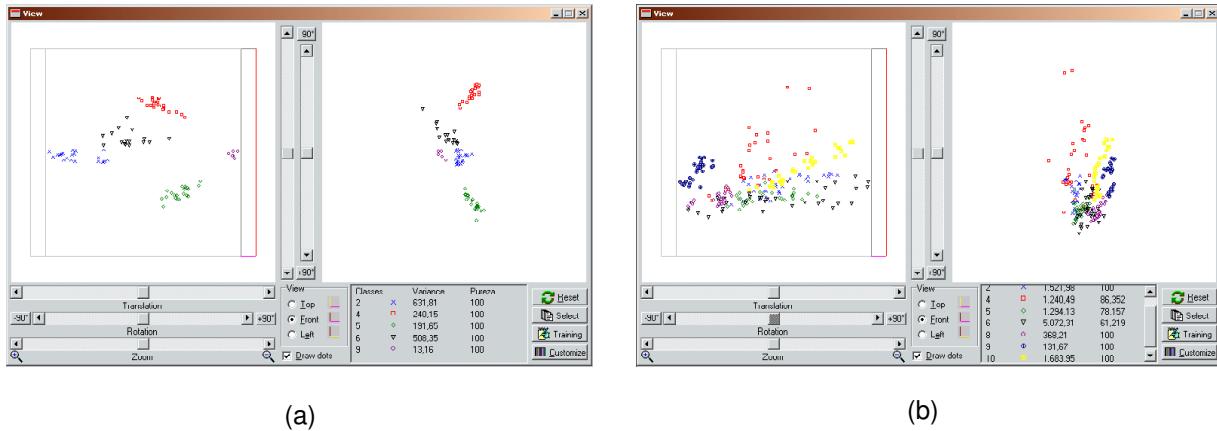


Figura 5.13: Visualização do extrator Gabor usando como métrica a distância Euclideana para (a) grupo de teste *TesteBásico* e (b) grupo de teste *TesteIntermediário*

Como pode ser visto na imagem 5.13 o primeiro grupo de teste (*GrupoBásico*) consegue ser melhor visualizado pelo fato de conter uma menor quantidade de imagens, o que não acontece no segundo grupo de teste (*Grupo Intermediário*), onde os vetores ficam embaralhados e os *clusters* apresentam interseção entre eles. Neste segundo caso considerou-se apenas os sete primeiros subgrupos, com a finalidade de visualizar a clusterização dos subgrupos e tendo em consideração que o grupo oito é considerado como ruído.

Na figura 5.14 apresenta-se o grafo de precisão e revocação, onde observa-se que este extrator com a distância Euclideana como métrica, não consegue ter um bom desempenho. Isto pode ser percebido na figura 5.15, onde mostra-se comparativamente o desempenho dos quatro extractores (*média*, *energia*, *entropia* e *gabor*) sendo a *entropia* a que consegue melhores resultados. A *entropia* dentro dos nove subespaços de alta freqüência indica o grau de suavidade, e as imagens médicas sob análise possuem extensas regiões de textura, por essa razão que a *entropia* conseguiu discriminar melhor as imagens.

5.6.4 Mudança na Medida de Distância

Até aqui só foi considerada a distância Euclideana como métrica de similaridade entre os vetores de características. No anexo B apresentam-se definições de outras medidas de distância. Mas o problema na extração de características é que cada extrator deve estar acompanhado de uma medida de distância, e isso vai depender do tipo de características que estão sendo extraídas.

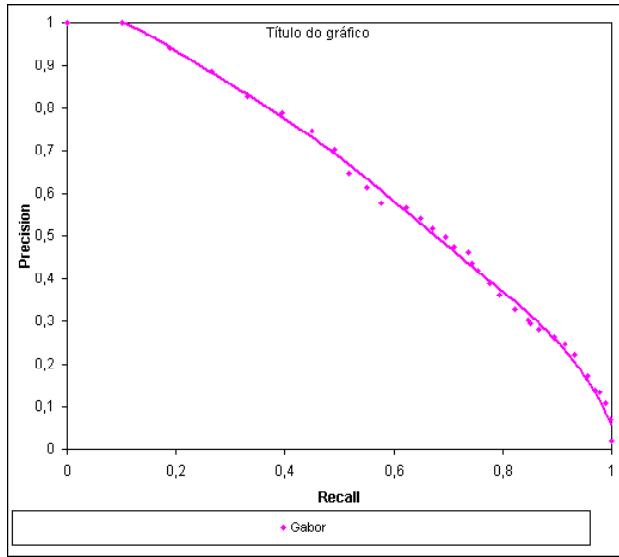


Figura 5.14: Grafo de precisão e revocação para o extrator de Gabor com a distância Euclideana como métrica, sobre o grupo de teste *GrupoIntermediário*.

Manjunath (Manjunath and Ma, 1996) define uma função distância (Euclideana normalizada) para o extrator baseado nas funções de Gabor a qual é definida a seguir. Considerando duas imagens padrões i e j , e sejam $\bar{f}^{(i)}$ e $\bar{f}^{(j)}$ que representam seus correspondentes vetores de características, então a distância entre os dois padrões no espaço de características é definida como:

$$d(i, j) = \sum_m \sum_n d_{mn}(i, j), \quad (5.5)$$

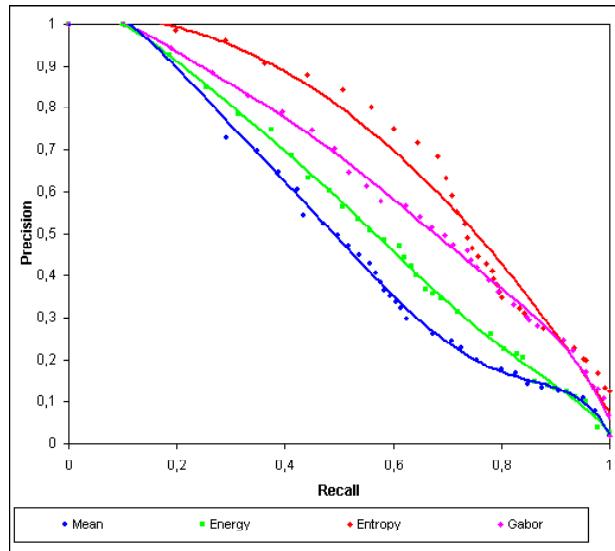


Figura 5.15: Grafo de precisão e revocação dos quatro extratores (*média*, *energia*, *entropia* e *gabor*), distância Euclideana e grupo de teste *GrupoIntermediário*.

onde

$$d_{mn}(i, j) = \left| \frac{\mu_{mn}^{(i)} - \mu_{mn}^{(j)}}{\alpha(\mu_{mn})} \right| + \left| \frac{\sigma_{mn}^{(i)} - \sigma_{mn}^{(j)}}{\alpha(\sigma_{mn})} \right| \quad (5.6)$$

sendo $\alpha(\mu_{mn})$ e $\alpha(\sigma_{mn})$ os desvios-padrões das respectivas características sobre o total da base de dados, e são usados para normalizar os componentes de características individuais.

Após fazer os testes de consulta com esta distância, sobre o conjunto de teste *GrupoIntermediário*, conseguiu-se os resultados que mostram-se na figura 5.16, onde o grafo tem um comportamento muito melhor dos que até agora foram vistos.

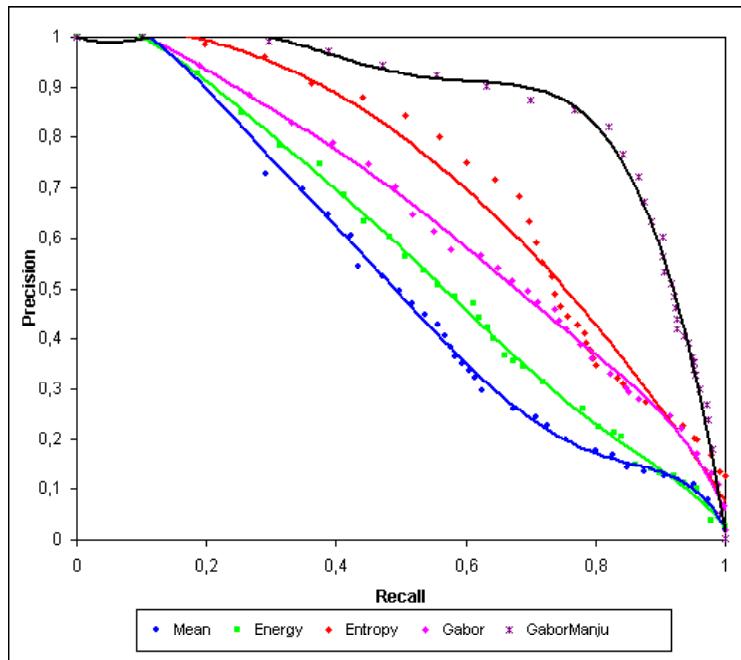


Figura 5.16: Visualização do extrator Gabor usando como métrica a distância definida por Manjunath (Manjunath and Ma, 1996) comparando com os outros extratores, grupo de teste *TesteIntermediário*

A desvantagem dessa medida de distância é que ela é uma métrica orientada ao espaço global dos vetores de características. Isto apresentará dificuldades no caso de existir inserções e eliminações de elementos do banco de dados, pelo fato que o cálculo de medidas estatísticas como desvio-padrão são feitos sobre o total dos elementos. Isto gera um novo problema nesta área de pesquisa: a procura de uma medida de distância, orientada ao vetor de características, que possa ser permitir inserções e eliminações de elementos, constituirão-se como problema a ser resolvido em futuros trabalhos de pesquisa.

5.7 Avaliação de Resultados no Sistema Protótipo

O teste final de um sistema CBIR só pode ser realizado através do uso efetivo do sistema por pessoal técnico especializado, sendo que a eficácia é medida pela precisão das consultas realizadas segundo o que é esperado pelo profissional médico. Para avaliar a utilidade da proposta, foi desenvolvido um protótipo capaz de realizar consultas por similaridade em um banco de dados de imagens reais (no apêndice C fornece-se da descrição do sistema e de algumas consultas de exemplo). Como foi detalhado previamente, o conjunto de imagens corresponde a grupos de exames tomográficos com imagens consecutivas.

5.7.1 Tempo de Geração de Características

O processo de geração de características consiste na transformação das imagens através de um filtro de *wavelet*, e análise dos espaços de *wavelet* para extrair e montar o vetor de características. A plataforma utilizada nos teste foi de um computador IBM/PC com processador Pentium III 866MHz e 128MB de memória RAM, operando sob sistema operacional Windows 2000.

O tempo gasto para calcular a transformada *wavelet* de Daubechies para o cálculo dos vetores de características de 1000 imagens de tamanho 256×256 , foi de 18,67 minutos. Ou seja, uma média de 1,12 segundos para aplicar a transformada *wavelet* e extrair as respectivas características.

Já com os filtros de Gabor o tempo foi maior, levando 10.17 minutos para gerar os vetores de 100 imagens de tamanho 256×256 , considerando 2 escalas e 6 orientações. Nesse caso o tempo médio ficou em 6.0 segundos para transformar as imagens e extrair as características com os filtros de Gabor. No caso de mudar para 4 escalas e 6 orientações, a geração dos vetores de características para 100 imagens de tamanho 256×256 levou 20.33 minutos, tendo uma média de 12.0 segundos por imagem.

A tabela 5.4 indica que o tempo de extração de características com *wavelet* de Gabor é proporcional ao número de características que estão sendo geradas, isto é, depende do número de escalas e orientações que estão sendo definidos.

Nro. Imagens	2-6 (escalas-orient.) (minutos)	4-6 (escalas-orient.) (minutos)	6-8 (escalas-orient.) (minutos)
01	00.10	00.20	00.38
10	01.17	02.33	03.90
30	02.98	06.05	12.22
100	10.17	20.33	39.00

Tabela 5.4: Tempo de geração de características baseadas em filtros de Gabor com distintos parâmetros (exemplo 2 – 6 = 2 escalas - 6 orientações).

5.7.2 Tempo de Resposta

O tempo de resposta a uma consulta é de crucial importância em um sistema CBIR, pois o usuário geralmente necessita da resposta imediatamente após a consulta. Devido à grande quantidade de informação médica que são gerados em centros hospitalares esse processo pode não ser rápido o suficiente. Para minimizar tal problema no cbPACS as imagens no banco de imagens são armazenadas juntamente com suas características previamente extraídas, o que é feito em horários ociosos da base. Durante o processamento da consulta do usuário efetuada pelo protótipo, serão extraídas as características apenas da imagem de consulta, gerando o seu vetor de características. Este vetor constitui-se no dado de entrada no processo de busca no banco de dados.

Para efeitos de teste, o sistema protótipo foi implementado utilizando indexação seqüencial, onde é calculado a distância do vetor consulta com cada um dos vetores do banco, para logo ser ordenados em forma ascendente de acordo às distâncias calculadas.

Esse processo levou uma média de 0.420 segundos para uma base de dados de 2815 imagens com 24 elementos em cada vetor de características. Para vetores compostos de 48 elementos cada, mantendo-se as mesmas 2815 imagens, o tempo médio de busca no banco de dados foi de 0.812 segundos. Ao se dobrar o tamanho do banco de imagens, isto é 5631, o tempo de resposta na consulta para vetores de 24 componentes foi de 0.865 e para vetores de dimensão 48 foi de 1.600 segundos. Na tabela 5.5 mostram-se esses números, o que gera um gráfico de comportamento constante para cada extrator, o que também acontece com o número de comparações.

Dimensão do Vetor	2815 imagens (segundos)	5631 imagens (segundos)
24	0.420	0.865
48	0.812	1.600

Tabela 5.5: Tempo médio de resposta de consulta no sistema protótipo usando indexação seqüencial junto à função Euclideana, para 5631 imagens.

O problema com este tipo de organização está no fato que o tempo aumenta à medida que o banco de dados cresce. Ao utilizar um método de acesso (estrutura de indexação) adequado, esse crescimento é diminuído. Outro fator importante com relação ao tempo foi a função de distância utilizada. No caso da distância Euclideana, o cálculo é feito comparando dois únicos vetores. No caso de mudar de função distância (Euclideana normalizada, por exemplo) esta requer do cálculo prévio de características globais do espaço de características (desvio padrão, media, etc.) o que incide no incremento do tempo de cálculo de distância entre vetores. Mostra-se na tabela 5.6 os resultados do tempo médio de resposta de acordo com o tipo de característica aplicado, em forma similar ao da tabela 5.5.

Os resultados apresentados nessa seção indicam a necessidade de se utilizar uma estrutura de indexação adequada para organizar as imagens e responder as consultas por similaridade.

Dessa forma, utilizou-se o método de acesso métrico *Slim-tree* (Traina et al., 2000), que foi desenvolvido no GBDI, para processar as consultas e os resultados deste teste são apresentados na próxima seção.

Dimensão do Vetor	2815 imagens	5631 imagens
24	1.820	2.510
48	2.220	4.570

Tabela 5.6: Resultados do tempo médio de resposta das consultas feitas no protótipo com uma outra distância (Euclideana normalizada), para 5631 imagens

5.8 Utilizando o Método de Acesso Métrico *Slim-tree*

Foi utilizada a primeira versão da *Slim-tree*, que trabalha em memória principal. Isto é, os objetos são indexados e a estrutura é armazenada em memória principal. Além disso, pode-se construir apenas uma árvore de cada vez, ou seja, se existem dois ou mais extratores suportados pelo sistema, somente uma árvore poderá ser montada a cada execução. Tal fato limita a utilização da estrutura em sistemas reais. Porém, a segunda versão da *Slim-tree*, já operando em memória secundária e sem tais restrições, deverá estar disponível em meados de 2003.

É importante analisar o comportamento das consultas ao se utilizar a *Slim-tree* para processá-las. Considerando o espaço de características gerados pelos filtros de Gabor com 2 escalas e 6 orientações (resultando em um vetor de características com 24 elementos), para uma base de 5631 imagens, o tempo de montagem da *Slim-tree* foi de 6.35 segundos. Já no caso dos filtros de Gabor com 4 escalas e 6 orientações (48 características por vetor), para o mesmo número de imagens (5631), o tempo de montagem foi de 13.46 segundos.

Embora a montagem da árvore tome alguns segundos, as respostas às consultas são muito rápidas, por exemplo na figura 5.17 mostram-se as médias dos tempos de resposta para diferentes consultas sobre os k -vizinhos mais próximos, onde Gabor-2-6 corresponde aos vetores gerados por filtros de Gabor com 2 escalas e 6 orientações; e Gabor-4-6 corresponde a 4 escalas e 6 orientações.

Grafo similar, mas com o cálculo do número de comparações para recuperar os k -vizinhos mais próximos na estrutura *Slim-tree* é apresentado na figura 5.18.

No caso do grafo da figura 5.18 a função correspondente à estrutura seqüencial, só foi considerada as comparações feitas apenas no processo de cálculo de distâncias, que é considerado como se for uma só comparação, o que é feito com o vetor de consulta e o total dos elementos do banco de imagens, neste caso de 5631. Não é considerado o tempo de ordenamento das distâncias (com o objetivo de uma melhor visualização do grafo), que no melhor caso é de $n \log n$ e no pior caso de n^2 (ordenamento *quicksort*).

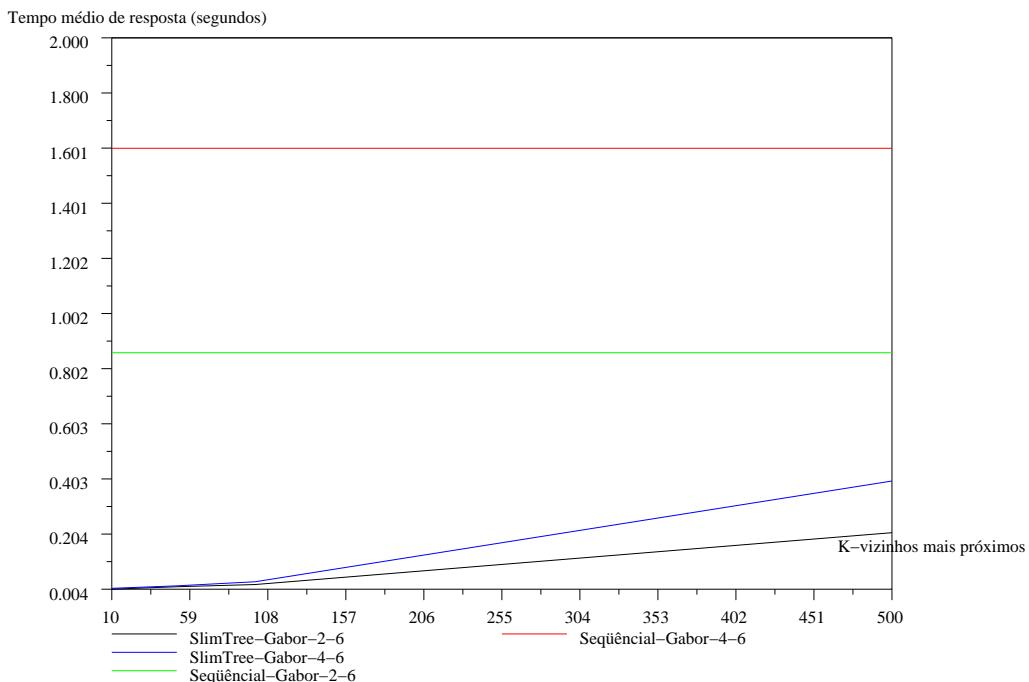


Figura 5.17: Médias dos tempos de resposta para diferentes consultas dos k -vizinhos mais próximos sobre a base de 5631 vetores de características utilizando a *Slim-tree* e busca seqüêncial.

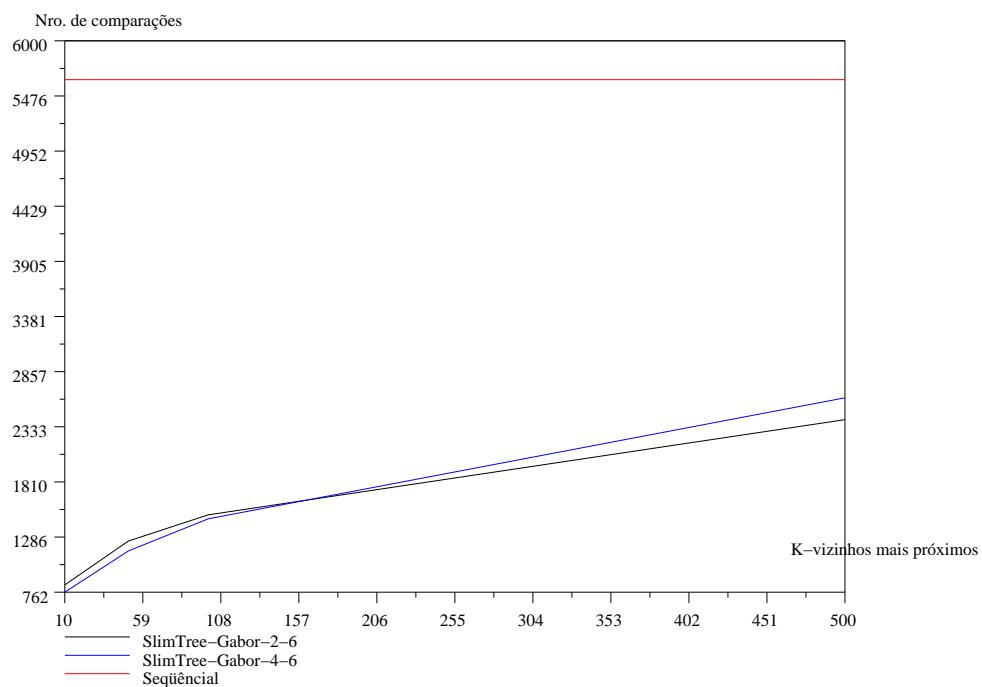


Figura 5.18: Grafo que mostra o número de comparações que são realizados na *Slim-tree* para recuperar k -vizinhos mais próximos, junto à estrutura seqüêncial (resposta constante) para um espaço composto de 5631 vetores de características.

5.9 Considerações Finais

Neste capítulo foram apresentados alguns resultados que mostram a capacidade da transformada *wavelet* para discriminar informação e, a partir desses novos espaços gerados, extrair características que possam ser utilizados para identificar imagens mediante um vetor de características.

Foram apresentados resultados procurando caracterizar as imagens em base a textura inata delas, para o que a *wavelet* de Gabor mostrou-se bastante adequada. Além de extrair as características desejadas deve-se utilizar funções de distância que consigam capturar a separação entre as imagens.

Finalmente, mostrou-se a grande importância que tem uma estrutura de indexação métrica, neste trabalho foi utilizada a *Slim-tree* sendo que os tempos de resposta para consultas por similaridade em uma estrutura métrica são muito menores do que com uma estrutura seqüencial, especialmente quando se dispõe de grandes quantidades de informação.

Conclusões

6.1 Considerações Finais

A pesquisa em recuperação de imagens por conteúdo ainda possui muitos desafios a serem atingidos. Um deles é a extração de características a partir de um domínio de imagens que permitam caracterizá-las mediante uma seqüência pequena de valores numéricos, permitindo assim reduzir a dimensionalidade dos objetos originais e indexá-los em uma estrutura métrica.

Neste trabalho, foram estudadas técnicas de extração de características em imagens através da transformada *wavelet*, sendo que o domínio de objetos correspondeu a imagens médicas. O enfoque foi sobre técnicas de caracterização de imagens aproveitando os espaços de *wavelet* gerados após aplicar os filtros de *wavelet*, para assim, mediante medidas estatísticas desses espaços, montar o vetor de características.

Esses estudos culminaram na realização de testes práticos com algumas técnicas de (a) caracterização baseadas na distribuição da cor e (b) caracterização baseada na textura das imagens. No primeiro teste utilizou-se as transformadas *wavelet* de Daubechies-2 (Haar) e de Daubechies-4 com medidas estatísticas para caracterizar o brilho (média), suavidade (entropia) e uniformidade (energia) das imagens. Para o segundo teste aplicou-se os filtros de Gabor, os quais permitem modelar o comportamento das células do cortex visual humano. As funções Gabor podem ser moduladas mediante mudanças na orientação e escala dos filtros, os quais geram subespaços transformados que contém micro-características que podem ser usadas para mapear a informação de textura inata da imagem.

Este trabalho contribuiu para a evolução do estado da arte de sistemas de recuperação de imagens médicas nos seguintes aspectos:

- Ampliou-se o domínio de extractores de características de imagens médicas a serem incluídas no sistema cbPACS, permitindo ter um novo domínio de espaço de características baseados na transformada *wavelet*.
- Considerando a necessidade de caracterizar imagens baseadas em sua textura, o trabalho foi estendido para aplicar outros filtros (além dos de Haar e Daubechies), encontrando-se que filtros de Gabor fornecem bons resultados.
- Criação de espaços de características que podem ser indexados em estruturas métricas e possibilitar o tratamento de consultas por similaridade diretamente.
- Mostrou-se o grande potencial das transformada *wavelet* não só na extração de características mas para outras aplicações, esperando-se que no futuro constituir-se-ão como uma ferramenta fundamental na análise e visualização de imagens médicas.
- Atenção especial merecem os filtros de Gabor, que junto com a análise de multiresolução constituem as *wavelets* de Gabor, que por suas especiais características fazem delas um dos melhores métodos na caracterização de textura.
- Sendo este o primeiro trabalho dentro do Grupo de Bases de Dados e Imagens do ICMC-USP em torno da extração de características aplicando transformada *wavelet*, envolvendo conceitos, definições, técnicas de caracterização, entre outros, passa a ter real valor como material de referência para estudos futuros dentro do grupo.

Além das contribuições colocadas acima é importante ressaltar que foi necessário desenvolver um sistema que permita (a) gerar o espaço de características através da transformação das imagens, (b) fazer consultas e (c) visualização dos resultados. Para isso foi aproveitado e estendido o conjunto de funções e objetos da biblioteca DicomLib (desenvolvida no GBDI) que processa imagens de formato DICOM e desenvolvido o sistema MultiWaveMed.

6.2 Linhas de Futuras Pesquisas

Os conceitos apresentados no presente estudo foram reunidos, ao final, objetivando a aplicação considerada, isto é, a caracterização de imagens médicas para serem indexadas em uma estrutura métrica. A pesquisa que subsidiou este trabalho possibilitou resposta a algumas questões, porém suscitou inúmeras outras, inerentes a toda investigação científica. A seguir, apresentam-se tópicos a serem tratados em trabalhos futuros:

- Durante os testes realizados, verificou-se que a mudança na função de distância tem uma alta incidência no resultado da consulta, mas muitas dessas funções distâncias estão baseadas na base toda, logo é interessante pesquisar sobre medidas de distância baseadas no vetor de características diretamente.
- Integrar ao sistema desenvolvido uma estrutura métrica de indexação, para tornar mais rápido o processo de recuperação de imagens.
- Montar vetores de características que caracterizem cada uma das regiões de textura da imagem, isto permitirá ter vetores de dimensão variável em um só espaço. Junto a isso, procurar por uma métrica que permita fazer as comparações e recuperar as imagens mais similares.
- Pesquisas devem ser realizadas para se obter uma representação que além da textura possa caracterizar a forma dos objetos. Isto beneficiaria no processamento das consultas de similaridade por regiões.

Essas são apenas algumas propostas de continuidade deste trabalho, mas muitas outras podem surgir após a leitura deste texto. Pode-se notar que os métodos propostos aqui não foram comparados com todos seus similares existentes, o que abre mais uma possibilidade de continuação deste trabalho.

Referências Bibliográficas

- Adoram, M. and Lew, M. (1999). IRUS: Image Retrieval Using Shape. *IEEE International Conference on Multimedia Computing and Systems*, 2:597–602.
- Albuz, E., Kocalar, E., and Khokhar, A. A. (2001). Scalable Color Image Indexing and Retrieval using Vector Wavelets. *IEEE Transaction on Knowledge and Data Engineering*, 13(5):851–861.
- A.Vailaya, Jain, A., and Zhang, H. (1998). On Image Classification: City Images vs. Landscapes". *Pattern Recognition*, 31(12):1921–1936.
- Baeza-Yates, R. and Ribeiro-Neto, B. (1999). *Modern Information Retrieval*. Addison Wesley.
- Baeza-Yates, R. A., Cunto, W., Manber, U., and Wu, S. (1994). Proximity Matching Using Fixed-Queries Trees. *Proceedings of the 5th Annual Symposium on Combinatorial Pattern Matching*, 1(807):198–212.
- Barioni, M. C., Botelho, E., Faloutsos, C., Razente, H., Traina, A. J., and Traina, Caetano, J. (2002). Data Visualization in RDBMS. In *IASTED Intl. Conference Information Systems and Databases (ISDB 2002)*, Tokyo, Japan.
- Beckmann, N., Kriegel, H.-P., Schneider, R., and Seeger, B. (1990). The R*-Tree: An Efficient and Robust Access Method for Points and Rectangles. In Garcia-Molina, H. and Jagadish, H. V., editors, *Proceedings of the 1990 ACM SIGMOD International Conference on Management of Data, Atlantic City, NJ, May 23-25, 1990*, pages 322–331. ACM Press.
- Bovik, A. C., Clark, M., and Geisler, W. S. (1990). Multichannel Texture Analysis Using Localized Spatial Filters. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(1):55–73.

- Bozkaya, T. and Özsoyoglu, Z. (1997). Distance-Based Indexing for High-Dimensional Metric Spaces. *Proc. ACM International Conference on Data Management (SIGMOD)*, pages 357–368.
- Bozkaya, T. and Özsoyoglu, Z. (1999). Indexing Large Metric Spaces for Similarity Search Queries. *ACM Transactions on Database System (TODS)*, 24(3):361–404.
- Bradshaw, B. (2000). Semantic Based Image Retrieval: A Probabilistic Approach. *ACM Multimedia*, pages 167–176.
- Brin, S. (1995). Near Neighbor Search in Large Metric Spaces. *Proc. International Conference on Very Large Databases (VLDB), Zurich, Switzerland*, pages 574–584.
- Bueno, J. M. (2002). *Suporte à Recuperação de Imagens Médicas Baseada em Conteúdo Através de Histogramas Métricos*. Tese de doutorado, Instituto de Ciências Matemáticas e de Computação - Universidade de São Paulo.
- Cesar, R. M. (1997). *Análise Multi-Escala de Formas Bidimensionais*. Tese de doutorado, Instituto de Física de São Carlos - Universidade de São Paulo.
- Chang, T. and Kuo, C. J. (1993). Texture Analysis and Classification with Tree-Structured Wavelet Transform. *IEEE Transaction on Image Processing*, 2(4):429–441.
- Chitre, Y. and Dhawan, A. P. (1999). M-Band Wavelet Discrimination of Natural Textures. *Pattern Recognition*, (32):773–789.
- Chiueh, T. C. (1994). Content-Based Image Indexing. *Proc. International Conference on Very Large Databases (VLDB), Santiago de Chile, Chile*, pages 582–593.
- Chávez, G. C. (2002). *Sistema Celular Evolutivo para Reconhecimento de Padrão Invariante*. Tese de mestrado, Instituto de Ciências Matemáticas e de Computação - Universidade de São Paulo.
- Ciaccia, P., Patella, M., and Zezula, P. (1997). M-tree: An Efficient Access Method for Similarity Search in Metric Spaces. *Proc. International Conference on Very Large Databases (VLDB), Athens, Greece*, pages 426–435.
- Cinque, L., Levialdi, S., Olsen, K., and Pellicanò, A. (1999). Color-Based Image Retrieval Using Spatial-Chromatic Histograms. *IEEE International Conference on Multimedia Computing and Systems*, 2:969–973.
- Coggins, J. and Jain, A. (1985). A Spatial Filtering Approach to Texture Analysis. *Pattern Recognition Letters*, 3:195–203.

- Costa, L., Kaye, B., and Montagnoli, C. (2000). Accurate Fractal Dimension Estimation using Exact Dilations. *IBERAMIA/SBIA 2000 - Workshop on Artificial Intelligence and Computer Vision*, pages 82–85.
- da Fontoura Costa, L. and Cesar, R. M. (2000). *Shape Analysis and Classification: Theory and Practice*. CRC Press.
- Daubechies, I. (1990). The Wavelet Transform, Time Frequency Localization, and Signal Analysis. *IEEE Transactions on Information Theory*, 36(5):961 – 1005.
- Daubechies, I. (1992). *Ten Lectures on Wavelets*, volume 61. CBMS-NSF Regional Conference Series in Applied Mathematics, Rutgers University and AT&T Bell Laboratories.
- Daugman, D. (1988). Complete Discrete 2-D Gabor Transforms by Neural Networks for Image Analysis and Compression. *IEEE Transactions on Acoustic, Speech, and Signal Processing*, 36(7):1169–1179.
- Daugman, D. (2002). How Iris Recognition Works. *In Proceedings of 2002 International Conference on Image Processing*, 1:33–36.
- Davis, G. M. and Nosratinia, A. (1998). Wavelet-Based Image Coding: An Overview. *Applied and Computational Control, Signals, and Circuits*, 1(44):205–269.
- de Lima, L. R. S., Laender, A. H. F., and Ribeiro-Neto, B. A. (1998). A Hierarchical Approach to the Automatic Categorization of Medical Documents. *Proceedings of the 1998 ACM CIKM International Conference on Information and Knowledge Management, Bethesda, Maryland, USA, November 3-7, 1998*, pages 132–139.
- Do, M. N., Ayer, S., and Vetterli, V. (1999). Invariant Image Retrieval using Wavelet Maxima Moment. *Proc. of 3th International Conference on Visual Information and Information System*.
- Feris, R. S., Krueger, V., and Jr., R. M. C. (2001). Efficient Real-Time Face Tracking in Wavelet Subspace. *In Proceedings of ICCV'2001 International Conference on Computer Vision, Workshop on REcognition, Analysis and Tracking of Faces and Gestures in Real-Time Systems, Vancouver, BC.*, pages 113–118.
- Flickner, M., Sawhney, H., and Niblack, W. (1995). Query by Image and Video Content: The QBIC System. *IEEE computer Magazine*, 28:23–32.
- Gabor, D. (1946). Theory of Communication. *J.IEE(London)*, 93, Part III(26):429–457.
- Gaede, V. and Günther, O. (1998). Multidimensional Access Methods. *ACM Computing Surveys*, 30(2):170–231.

- Gomez, J., Velho, L., and Goldenstein, S. (1997). *Wavelets: Teoria, Software e Aplicações*. 21 Colóquio Brasileiro de Matemática. Instituto de Matemática Pura e Aplicada, Rio Janeiro - Brasil.
- Gonzales, R. and Woods, R. (1993). *Digital Image Processing*. Addison Wesley.
- Grossmann, A. and Morlet, J. (1984). Decomposition of Hardy Function into Square Integrable Wavelets of Constant Shape. *SIAM, Journal of Mathematical Analysis*, 15(4):723–736.
- Guttman, A. (1984). R-Trees: A Dynamic Index Structure for Spatial Searching. *Proceeding of the International Conference on Data Management(ACM-SIGMOD)*, Boston, MA, pages 47–57.
- He, C., Dong, J., Sheng, Y. F., and Ahalt, S. C. (2001). Object Tracking Using the Gabor Wavelet Transform and the Golden Section Algorithm. *Proceeding of the 2001 IEEE International Conference on Robotic and Automation*, Seoul, Korea, 1:1671–1676.
- Hellerstein, J. M., Naughton, J. F., and Pfeffer, A. (1995). Generalized Search Trees for Database Systems. In Dayal, U., Gray, P. M. D., and Nishio, S., editors, *Proc. 21st Int. Conf. Very Large Data Bases, VLDB*, pages 562–573. Morgan Kaufmann.
- Hinneburg, A., Keim, D. A., and Wawryniuk, M. (1999). HD-Eye: Visual Mining of High-Dimensional Data. *IEEE Computer Graphics and Applications*, 19(5):22–31.
- Hu, M.-K. (1962). Visual Pattern Recognition by Moment Invariants. *IRE Transaction on Information Theory*, IT-8:179–187.
- Hubbard, B. B. (1998). *The World According to Wavelets: The History of a Mathematical Technique in the Making*. A K Peters, Wellesley, Massachusetts, 2nd. edition.
- Jacobs, C. E., Finkelstein, A., and Salesin, D. H. (1995). Fast Multiresolution Image Querying. *Computer Graphics*, 29(Annual Conference Series):277–286.
- Jain, A. and Farrokhnia, F. (1991). Unsupervised Texture Segmentation using Gabor Filters. *Pattern Recognition*, 24:1167–1186.
- Jain, A. and Healey, G. (1998). A Multiscale Representation Including Opponent Color Features for Texture Recognition. *IEEE Transactions in Image Processing*, 7(1):124–128.
- Jain, A., Ross, A., and Prabhakar, S. (2001). Fingerprint Matching using Minutiae and Texture Features. Technical Report MSU-CSE-01-17, Department of Computer Science, Michigan State University, East Lansing, Michigan.
- Jain, A. K. and Vailaya, A. (1998). Shape-Based Retrieval: A Case Study with Trademark Image Databases. *Pattern Recognition*, 21:1369–1390.

- Korn, F., Sidiropoulos, N., Faloutsos, C., Siegel, E., and Protopapas, Z. (1996). Fast Nearest Neighbor Search in Medical Image Databases. *The VLDB Journal*, pages 215–226.
- Krueger, V. and Sommer, G. (1999). Affine Real-Time Face Tracking Using a Wavelet Network. In *Proceedings of ICCV'99 Workshop Recognition, Analysis, and Tracking of Faces and Gestures in Real-Time Systems, Corfu, Greece*.
- Lee, C.-J., Wang, S.-D., and Wu, K.-P. (2001). Fingerprint Recognition Using Principal Gabor Basis Function. In *Proceedings of 2001 International Symposium on Intelligent Multimedia, Video and Speech Processing - Hong Kong*, pages 393–396.
- Lee, T. S. (1996). Image Representation Using 2D Gabor Wavelets. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(10):959–971.
- Loew, M. H. (2000). *Feature Extraction*, chapter 5. SPIE, Bellingham, WA, m.sonka and j. michael fitzpatrick edition.
- Ma, W. and Manjunath, B. (1995). Image Indexing Using a Texture Dictionary. in *Proc. SPIE Conference on Image Storage and Archiving System*, 2606:288–298.
- Malik, J. and Perona, P. (1990). Preattentive Texture Discrimination with early Vision Mechanisms. *Opt. Soc. Am. Series*, A(7):923–932.
- Mallat, S. G. (1988). *Multiresolution Representation and Wavelets*. Ph.d. thesis, University of Pennsylvania.
- Mallat, S. G. (1989). A Theory for Multiresolution Signal Decomposition: The Wavelet Representation. *IEEE Transaction on Pattern analysis and Machine Intelligence*, 11(7):674–693.
- Mallat, S. G. (1999). *A Wavelet Tour of Signal Processing*. Academic Press, 2nd. edition.
- Manjunath, B. and Ma, W. (1996). Texture Features for Browsing and Retrieval of Image Data. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 18:837–842.
- Mao, J. and Jain, A. (1992). Texture Classification and Segmentation Using Multiresolution Simultaneous Autoregressive Models. *Pattern Recognition Journal*, 25(2):173–188.
- Marsh, A. (1997). Euromed - The Creation of a Telemedical Information Society. *Tenth IEEE Symposium on Computer-Based Medical System*, pages 86–91.
- Meyer, Y. (1993). *Wavelets: Algorithms and Applications*. Society for Industrial and Applied Mathematics, Philadelphia, first edition.
- Misiti, M., Oppenheim, G., Poggi, J.-M., and Misiti, Y. (2001). Wavelet Toolbox Documentation.

- Olshausen, B. A. (2002). Principles of Image Representation in Visual Cortex. *To appear in The Visual Neurosciences, L.M. Chalupa, J.S. Werner, MIT Press.*
- Papadias, D., Sellis, T., Theodoridis, Y., and Egenhofer, M. J. (1995). Topological Relations in the World of Minimum Bounding Rectangles: A Study with R-trees. In *Proc. ACM Int. Conf. on Data Management (SIGMOD), San José, CA*, pages 92–103.
- Pass, G. and Zabih, R. (1999). Histogram Refinement for Content-Based Image Retrieval. *IEEE Workshop on Applications on Digital Libraries*, 7(3):234–240.
- Peters, T. M. and Williams, J. (1998). *The Fourier Transform in Biomedical Engineering*. Birkhäuser.
- Petkov, N. and Kruizinga, P. (1997). Computational Models of Visual Neurons Specialised in the Detection of Periodic and Aperiodic Oriented Visual Stimuli: Bar and Grating Cells. *Biological Cybernetics*, 76(2):83–96.
- Picard, R. and T.P.Minka (1995). Vision Texture for Anotation. *Multimedia Systems: Special Issue on Content-Based Retrieval*, 3:3–14.
- Rao, R. P., Olshausen, B. A., and Lewicki, M. S. (2002). *Probabilistic Models of the Brain: Perception and Neural Function*. MIT Press.
- Rodrigues, J. F. (2003). Desenvolvimento de um Framework para Análise Visual de Informações Suportando Data Mining em um Sistema de Arquivamento de Imagens Médicas. Dissertação de Mestrado - ICMC-USP.
- Rubner, Y. and Tomasi, C. (2000). *Perceptual Metrics for Image Database Navigation*, volume 594 of *Series in Engineering and Computer Science*. Kluwer Academic Publisher, Boston.
- Sabino, D., Nakamura, E., Costa, L., Calado, R., and Zago, M. (2002). Chromatin Texture Characterization using Multiscale Fractal Dimension. *14th International Conference on Digital Signal Processing*, 2:529–533.
- Sellis, T., Roussopoulos, N., and Faloutsos, C. (1987). The R+-Tree: A Dynamic Index for Multi-Dimensional Objects. *Proceeding of the International Conference of Very Large Data Bases, Brighton, England*, pages 507–518.
- Sheikholeslami, G., Chatterjee, S., and Zhang, A. (1998). WaveCluster: A Wavelet-Based Clustering Approach for Multidimensional Data in Very Large Databases. *Conference on Very Large DataBase - VLDB'98*.
- Smeulders, A. W., Worring, M., Santini, S., Gupta, A., and Jain, R. (2000). Content-Based Image Retrieval at the End of the Early Years. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(12):1349–1380.

- Sonka, M., Hlavac, V., and Boyle, R. (1999). *Image Processing, Analysis, and Machine Vision*. Pacific Grove, CA: PWS Pub, 2nd. edition.
- Stollnitz, E. J., Deroose, T. D., and Salesin, D. H. (1996). *Wavelets for Computer Graphics, Theory and Applications*. Morgan Kaufmann Publisher, Inc, San Francisco, California, 1 edition.
- Swain, M. and Ballard, D. (1991). Color Indexing. *International Journal of Computer Vision*, 7(1):11–32.
- Theodoridis, S. (1999). *Pattern Recognition*. Academic Press, Greece.
- Traina, A. J. (2001). *Supporte à Visualização de Consultas por Similaridade em Imagens Médicas através de Estruturas de Indexação Métrica*. Tese de livre-docente em computação, Instituto de Ciências Matemática e de Computação - Universidade de São Paulo, São Paulo - Brasil.
- Traina, C., Traina, A., Seeger, B., and Faloutsos, C. (2000). Slim-Trees: High Performance Metric Trees Minimizing Overlap Between Nodes. *Proc. International conference on Extending Database Technology, Konstanz, Germany*, pages 51–65.
- Tuceryan, M. and Jain, A. (1993). Texture Analysis. in *Handbook of Pattern Recognition and Computer Vision (C.H. Chen, L.F. Pau, and P.S.P Wang, eds)*, pages 235–276. World Scientific Publishing Company.
- Turner, M. R. (1986). Texture Discrimination by Gabor Functions. *Biological Cybernetics*, 55(2-3):71–82.
- Unser, M. and Aldroubi, A. (1996). A Review of Wavelets in Biomedical Applications. *Proceeding of the IEEE*, 84(4):626–638.
- Vailaya, A. (2000). *Semantic Classification in Image Databases*. Phd. dissertation, Michigan State University.
- Wang, J. Z. (2000). *Semantics-Sensitive Integrated Matching for Picture Libraries and Biomedical Image Databases*. Phd. dissertation, Departament of Biomedical Informatics of Stanford University.
- Wang, J. Z. (2001). Methodological Review - Wavelets and Imaging Informatics: A Review of the Literature. *Journal of Biomedical Informatics*, pages 129–141. <http://www.idealibrary.com>.
- Wang, J. Z., Li, J., and Wiederhold, G. (2001). SIMPLICITY: Semantic-Sensitive Integrated Matching for Picture Libraries. *IEEE Transaction on Pattern Analysis and MACHine Intelligence*, 23(9):947–963.

- Weiss, G. (2002). Welcome to the (Almost) Digital Hospital. *IEEE Spectrum*, 39(3):44–49.
- Wilson, D. and Martinez, T. (1997). Improved Heteregeneous Distance Functions. *Journal of Artificial Inteligence Research*, (6):1–34.
- Xiao, Y., Wang, Y., and Felleman, D. J. (2003). A Spatially Organized Representation of Colour in Macaque Cortical Area V2. *Nature*, 421:535–539.
- Yong Rui, Thomas S. Huang, S.-F. C. (1999). Image Retrieval: Current Techniques, Promising Directions and Open Issues. *The NECI Scientific Literature Digital Library*.
- You, J. and Bhattacharya, P. (2000). Dynamic Shape Retrieval by Hierarchical Curve Mat- ching, Snakes and Data Mining. *15th International Conference on Pattern Recognition*, 1:1035–1038.

Revisão de Álgebra Linear

A.1 Espaço Vetorial ou Linear

Um espaço vetorial sobre os reais é um conjunto V de elementos, onde:

- Para todo $a, b \in \mathbb{R}$ e para todo $u, v \in V$, $au + bv \in V$.
- Existe um único elemento $0 \in V$ tal que:
 - para todo $u \in V$, $0u = 0$;
 - para todo $u \in V$, $0 + u = u$.

Os elementos de um espaço vetorial V são chamados vetores e o elemento zero é chamado de vetor nulo. Os vetores podem ser vetores genéricos, ou eles podem ser funções, como é o caso em *wavelets*.

A.2 Bases e Dimensões

Um conjunto de vetores $\{u_1, u_2, \dots\}$ em um espaço vetorial V são ditos linearmente independentes se: $c_1u_1 + c_2u_2 + \dots = 0$ se e somente se $c_1 = c_2 = \dots = 0$.

Um conjunto $\{u_1, u_2, \dots\} \in V$ de vetores linearmente independentes são uma base para V se todo $v \in V$ pode ser escrito como:

$$v = \sum_i c_i u_i$$

para quaisquer números reais c_1, c_2, \dots

Se uma base para V tem um número finito de elementos $\{u_1, u_2, \dots, u_m\}$ então V é de dimensão finita e sua dimensão é m , caso contrário V é de dimensão infinita.

Exemplo:

\mathbb{R}^3 é um espaço tridimensional.

Os vetores: $e^1 = (1, 0, 0)$, $e^2 = (0, 1, 0)$ e $e^3 = (0, 0, 1)$ formam uma base para \mathbb{R}^3 , ou seja, qualquer elemento de \mathbb{R}^3 pode ser escrito em função de e^1 , e^2 e e^3 .

A.3 Corpo

Um conjunto $F \neq \emptyset$ é um corpo comutativo se existem aplicações de $F \times F \rightarrow F : (\alpha, \beta) \mapsto \alpha + \beta$ e de $F \times F \rightarrow F : (\alpha, \beta) \mapsto \alpha\beta$, denominadas adição e multiplicação, respectivamente. Ou seja, um corpo é um conjunto munido de algumas operações sobre seus elementos, que se comportam como a adição, subtração, multiplicação e divisão, usuais de números, no sentido de que elas obedecem certas propriedades:

- $\alpha + (\beta + \gamma) = (\alpha + \beta) + \gamma$
- $\alpha + \beta = \beta + \alpha$
- $\exists 0 \in F$ tal que $\alpha + 0 = \alpha$
- $\exists -\alpha \in V$ para o qual $\alpha + (-\alpha) = 0$
- $\alpha\beta = \beta\alpha$
- $(\alpha\beta)\gamma = \alpha(\beta\gamma)$
- $\exists 1 \in F \setminus \{0\}$ tal que $\alpha 1 = \alpha$
- $\exists \alpha^{-1} \in F \setminus \{0\}$ tal que $\alpha\alpha^{-1} = 1$
- $\alpha(\beta + \gamma) = \alpha\beta + \alpha\gamma$

A.4 Produto Interno e Ortogonalidade

Seja K o corpo dos números reais ou complexos e V um espaço vetorial sobre K . Um produto interno sobre V é um mapeamento $\langle , \rangle : V \times V \rightarrow K : (u, v) \mapsto \langle u, v \rangle$. Denomina-se o

par $(V, \langle \cdot, \cdot \rangle)$ de espaço com produto interno (ou espaço pré-Hilbert) sobre K se as seguintes propriedades são observadas.

Para todo $u, v, w \in V$ e $\alpha \in K$ deve-se ter:

- $\langle u + v, w \rangle = \langle u, w \rangle + \langle v, w \rangle$
- $\langle \alpha u, w \rangle = \alpha \langle u, w \rangle$
- $\langle u, v \rangle = \overline{\langle v, u \rangle}$
- $\langle u, u \rangle \geq 0$
- $\langle u, u \rangle = 0 \Rightarrow u = 0$

$\langle \cdot, \cdot \rangle$ é denominado de produto interno.

Um dos usos mais importantes do produto interno é para formalizar a idéia de ortogonalidade. Dois vetores u e v são ditos ortogonais se $\langle u, v \rangle = 0$.

Uma base ortogonal é uma base que consiste de vetores ortogonais entre si.

A.5 Normas e Normalização

Uma norma é uma função que mede o tamanho de um vetor. Em um espaço vetorial de dimensão finita, freqüentemente usa-se a norma-2

$$\| u \|_2 := \langle u, u \rangle^{1/2}.$$

Um vetor u com $\| u \| = 1$ é dito normalizado.

Se temos uma base ortogonal composta de vetores normalizados, a base é chamada ortonormal.

Exemplo:

Os vetores e^1, e^2 e e^3 do exemplo anterior formam uma base ortonormal para o espaço \mathbb{R}^3 .

Medidas de Distância

Funções de distância são utilizadas em várias áreas, incluindo aprendizagem baseada em instâncias, redes neurais, estatísticas, reconhecimento de padrões e psicologia cognitiva. Muitos sistemas inteligentes dependem da eficiência de uma função para cálculo de distância entre dois vetores. Uma variedade de funções de distância está disponível para tais usos, incluindo as métricas de distância Minkowsky, Mahalanobis, Camberra, Chebychev, Quadrática, Correlação e Chi-quadrado. As funções mais comuns são definidas a seguir (Wilson and Martinez, 1997):

- Minkowsky:

$$d(x, y) = \left(\sum_i^m |x_i - y_i|^r \right)^{\frac{1}{r}} \quad (\text{B.1})$$

- Euclidean:

$$d(x, y) = \sqrt{\sum_{i=1}^m (x_i - y_i)^2} \quad (\text{B.2})$$

- Manhattan/city-block:

$$d(x, y) = \sum_{i=1}^m |x_i - y_i| \quad (\text{B.3})$$

- Camberra:

$$d(x, y) = \sum_{i=1}^m \frac{|x_i - y_i|}{|x_i + y_i|} \quad (\text{B.4})$$

- Chebychev:

$$d(x, y) = \max_{i=1}^m |x_i - y_i| \quad (\text{B.5})$$

- Quadratic:

$$d(x, y) = (x - y)^T Q (x - y) = \sum_{j=1}^m \left(\sum_{i=1}^m (x_i - y_i) q_{qi} \right) (x_j - y_j) \quad (\text{B.6})$$

Q é uma matriz de tamanho $m \times n$ finita, positiva e específica ao problema.

- Mahalanobis:

$$d(x, y) = [det V]^{\frac{1}{m}} (x - y)^T V^{-1} (x - y) \quad (\text{B.7})$$

V é a matriz de covariância de $A_1 \dots A_m$, e A_j é o vetor de valores para o atributo j que aparece nas instâncias do conjunto de treinamento $1 \dots n$.

- Correlação:

$$d(x, y) = \frac{\sum_{i=1}^m (x_i - \bar{x}_i)(y_i - \bar{y}_i)}{\sqrt{\sum_{i=1}^m (x_i - \bar{x}_i)^2 \sum_{i=1}^m (y_i - \bar{y}_i)^2}} \quad (\text{B.8})$$

$\bar{x}_i = \bar{y}_i$ é o valor médio para o atributo i que aparece no conjunto de treinamento.

- Chi-quadrado:

$$d(x, y) = \sum_{i=1}^m \frac{1}{soma_i} \left(\frac{x_i}{tam_x} - \frac{y_i}{tam_y} \right)^2 \quad (\text{B.9})$$

$soma_i$ é a soma de todos os valores para o atributo i que aparece no conjunto de treinamento, e tam_x é a soma de todos os valores no vetor x .

- Correlação de Posição de Kendall:

$$d(x, y) = 1 - \frac{2}{m(n-1)} \sum_{i=1}^m \sum_{j=1}^{i-1} sinal(x_i - x_j) sinal(y_i - y_j) \quad (\text{B.10})$$

$sinal(x) = -1, 0 ou 1$ se $x < 0$, $x = 0$, ou $x > 0$, respectivamente.

Descrição do Protótipo de Sistema de Consulta

O sistema está baseado na abordagem de multiresolução de *wavelet* para extrair características e recuperar imagens por conteúdo. O processo central do sistema corresponde ao processo de transformação de imagens em um vetor de características. A figura C.1 indica o esquema geral do sistema de consulta por similaridade.

C.1 Transformação de Imagens

Esse é o processo central do sistema, onde uma imagem (ou conjunto de imagens) é transformada em um conjunto de valores numéricos, nomeado “*vetor de características*”. A transformação é feita aplicando-se uma convolução entre a imagem de entrada e o filtro de *wavelet* (Daubechies, Gabor), tendo em consideração a abordagem de multiresolução. Após transformada a imagem são aplicadas algumas medidas estatísticas que tentam identificar algumas características da imagem como o brilho (média), contraste (variância), complexidade e aleatoriedade (entropia e variância). Esse conjunto de valores de características constituirão o vetor de características.

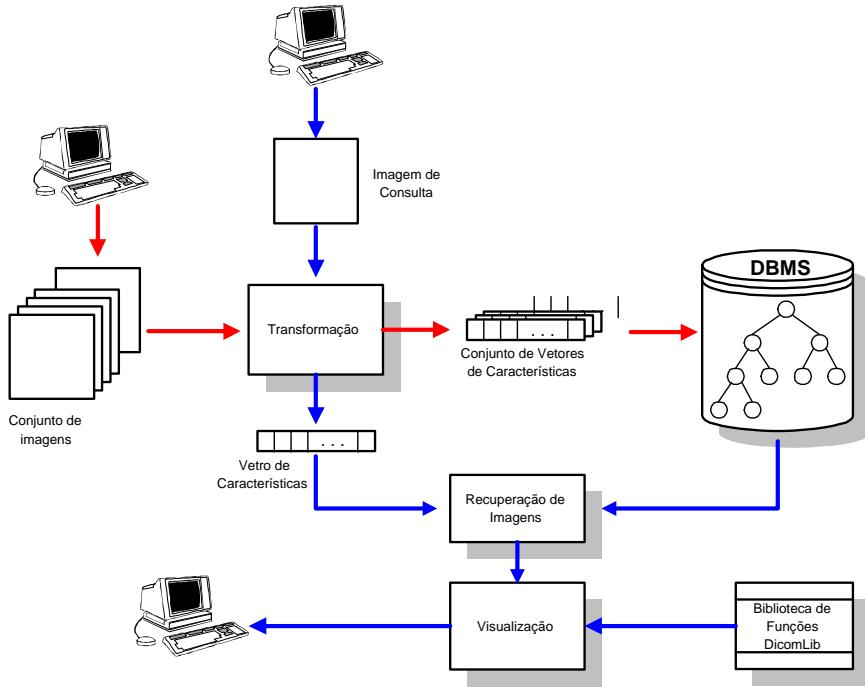


Figura C.1: Esquema do sistema de consulta por similaridade

C.2 Geração do Espaço de Características

Esse módulo considera um conjunto de imagens mais um filtro de *wavelets* (Daubechies, Gabor) para gerar os vetores de características através do processo de transformação. O resultado é armazenado em uma base de dados, onde cada imagem (nome e localização) tem seu correspondente vetor de características, o número de valores de cada vetor (dimensão) é uma função do filtro *wavelets* e dos parâmetros do filtro.

C.3 Recuperação de Imagens

Esse processo requer uma imagem de consulta, a qual é transformada em um vetor de características através do processo de transformação. Consequentemente, esse novo vetor gerado é procurado no banco de dados e, dependendo do tipo de consulta (vizinhos mais próximos ou consulta por abrangência) são recuperadas as imagens mais similares à imagem de consulta, baseados em uma métrica (função distância).

C.4 Sistema de Indexação de dados

No momento o sistema de indexação de dados é seqüencial, mas o sistema de recuperação possui a capacidade de suportar um outro sistema de indexação como a estrutura métrica *Slim-tree*.

C.5 Visualização de Consultas

O módulo de visualização de consultas foi desenvolvido em C++Builder, e parte importante desse está constituído pela biblioteca de funções *DicomLib* a qual é um conjunto de objetos e funções que trabalham com imagens médicas (formato Dicom) que fornecem um adequado conjunto de funções para o processamento desse formato de imagens. Esta biblioteca foi desenvolvida no nosso laboratório GBDI.

A seguir apresentam-se exemplos de consultas por similaridade do sistema os que estão baseados nos distintos extractores desenvolvidos neste trabalho. A primeira imagem mostra os componentes principais da tela inicial do sistema:

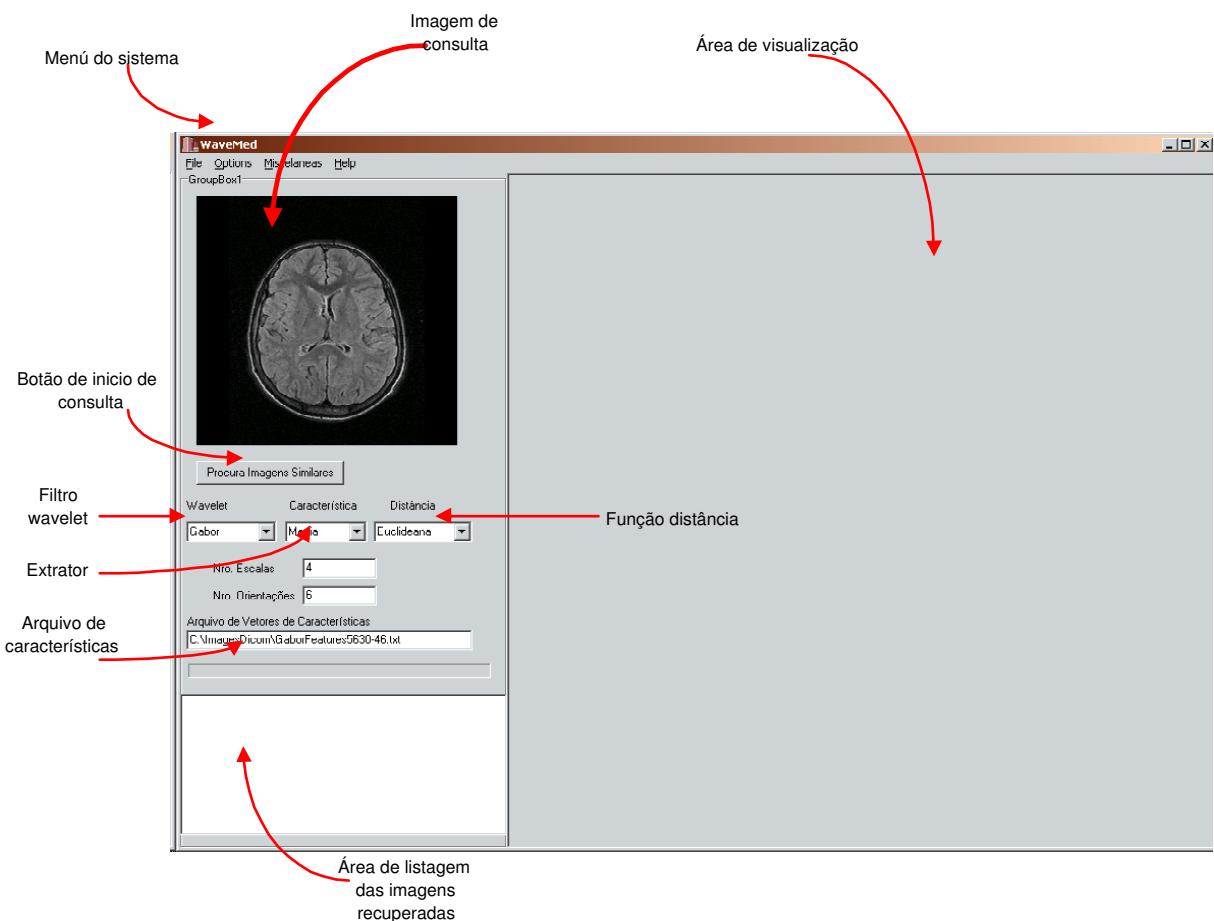


Figura C.2: Tela principal do sistema

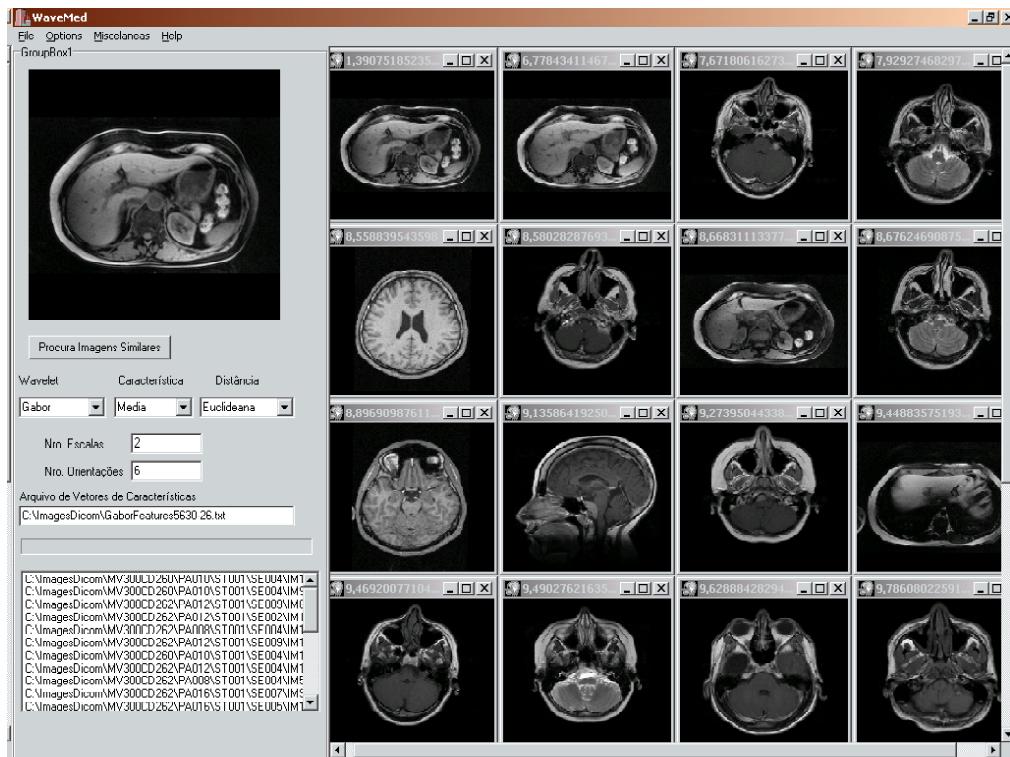


Figura C.3: Consulta com Gabor-2-6 utilizando distância Euclideana

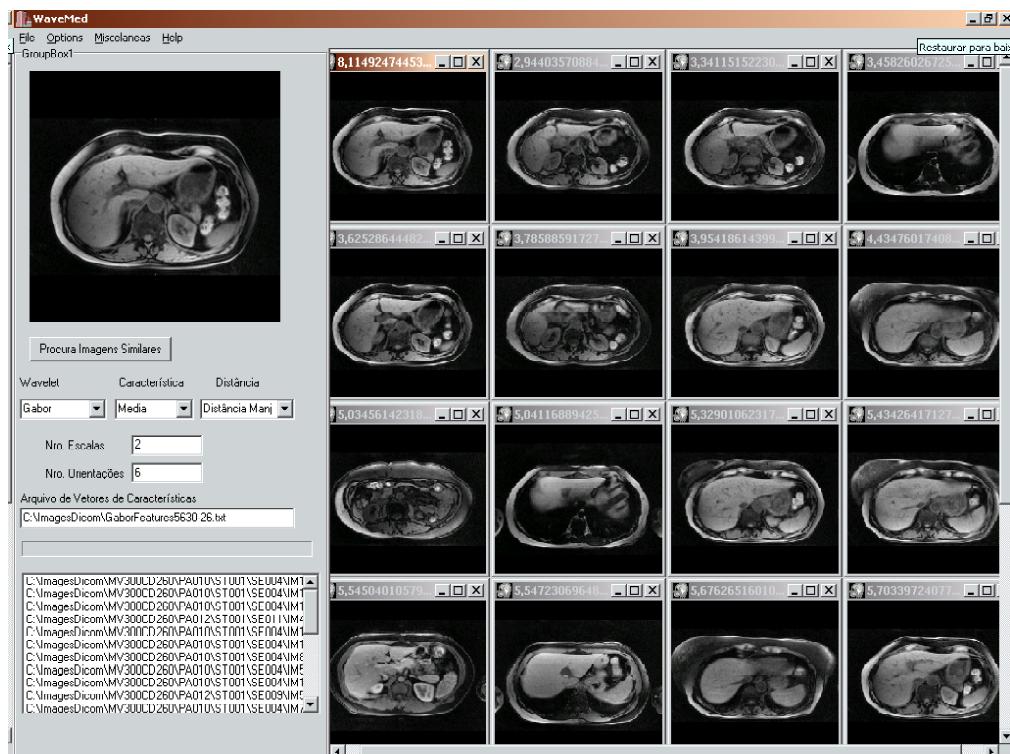


Figura C.4: Consulta com Gabor-2-6 utilizando distância Euclideana Normalizada

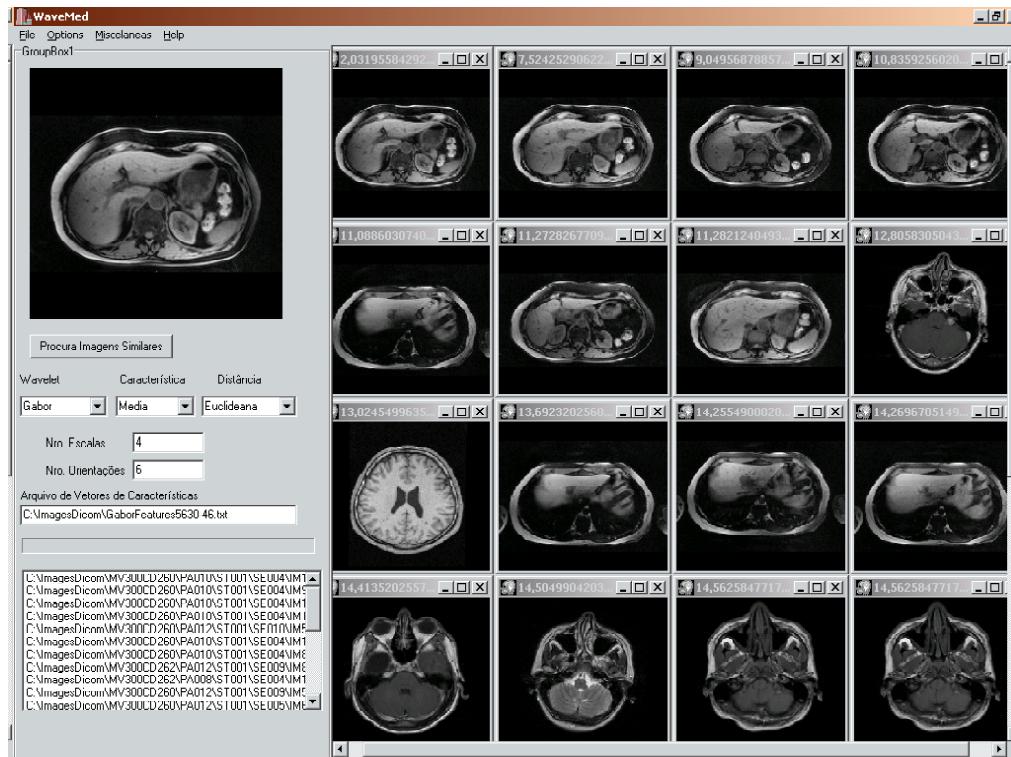


Figura C.5: Consulta com Gabor-4-6 utilizando distância Euclideana



Figura C.6: Consulta com Gabor-4-6 utilizando distância Euclideana Normalizada



Figura C.7: Consulta com Gabor-2-6 utilizando distância Euclideana Normalizada, procurando uma imagem que não pertence ao banco de imagens

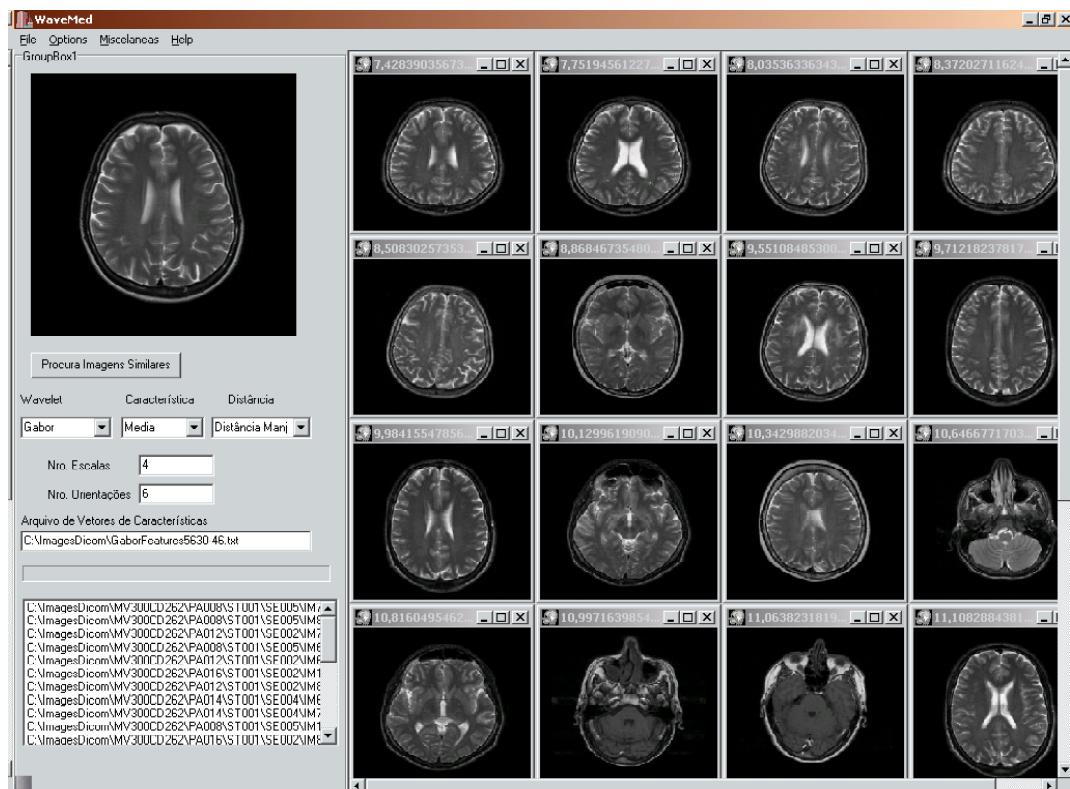


Figura C.8: Consulta com Gabor-4-6 utilizando distância Euclideana Normalizada, procurando uma imagem que não pertence ao banco de imagens



Figura C.9: Consulta com Gabor-4-6 utilizando distância Euclideana Normalizada de uma imagem aleatória

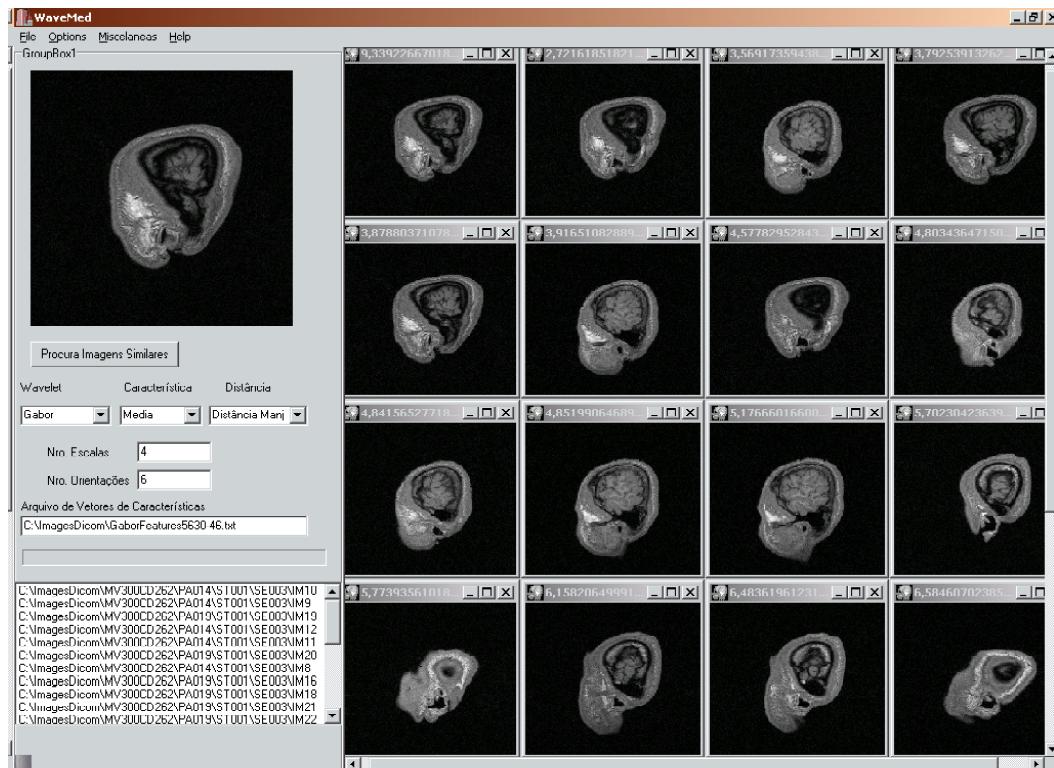


Figura C.10: Consulta com Gabor-4-6 utilizando distância Euclideana Normalizada de uma imagem aleatória