

Discriminação de Texturas Pela Transformada Wavelet

Aylton Pagamisse

Prof. Dr. Luciano Vieira Dutra
Orientador

Prof^a. Dr^a. Maria Cristina C. Cunha
Co-Orientadora

Tese apresentada ao Instituto de Matemática, Estatística e Computação Científica, UNICAMP, como requisito parcial para a obtenção do Título de Doutor em Matemática Aplicada.

Aos meus pais Armando Pagamisse e Rosa Dadalto Pagamisse.

Agradecimentos

Ao meu orientador Luciano Vieira Dutra, pela orientação, pelas idéias,, por me mostrar um mundo novo e pela amizade nesses anos.

À Prof^a. Maria Cristina C. Cunha, orientadora do programa de doutorado, que me deu a oportunidade de ser aluno da Unicamp, pela amizade e pelos muitos ensinamentos.

Aos colegas do Departamento de Matemática, Estatística e Computação da FCT - UNESP.

Aos colegas de convivência nesse período em Campinas: José Roberto, Socorro, Marcio Rodolfo, Cristiane, Paulo, Vanderlei, Márcio, Galo, Lourdes, Dino e Rosângela.

Aos professores do DMA – IMECC – UNICAMP e aos funcionários da Seção de Pós-Graduação do IMECC – UNICAMP, pela ajuda e apoio constante.

À Capes/PICD, pelo apoio financeiro.

A todos que direta ou indiretamente colaboraram para a realização deste trabalho.

*Aí rapaz
Dediquei toda minha vida ao mundo das imagens
artificiais dos vídeos, das telinhas, dos telões
vetezei dos comerciais mais lisérgicos
aos mais belos bombardeios aéreos
Por isso eu te digo rapaz
pra cada beijo e facada existe uma coisa pesquisada
E o mais vagabundo ferro de passar
Tem a ver com uma pesquisa militar
Quantas vezes o mundo é registrado, catalogado todos os dias?*

*Por isso eu te digo rapaz
Que diante das imagens é preciso ter
flipperglamour
E relaxar a razão de todas as coisas*

*Trecho da Música "Chinesa Videomaker"
Fausto Fawcett & Carlos Laufer
1984*

Resumo

Em imagens de sensoriamento remoto e de outras aplicações, tais como análise de imagens médicas, inspeção industrial, etc, é necessária, em muitos casos, a discriminação de texturas. Intuitivamente, texturas estão relacionadas a variações tonais na imagem, sendo propriedade de uma região e não de um único ponto. Essas variações podem ser vistas como descontinuidades locais e, em imagens, estão, em geral, associadas a bordas.

A transformada Wavelet fornece uma representação completa e estável de sinais utilizando bordas em multiescala. Nesse trabalho, desenvolve-se uma representação Wavelet invariante por translações, com quatro direções e, utilizando essa representação, desenvolve-se um método que agrega e combina as respostas da energia da transformada Wavelet nas várias escalas e direções para obter a discriminação de texturas.

Esse método tem baixo custo computacional, pois obtém seus parâmetros a partir de amostras de texturas, e a implementação discreta da decomposição Wavelet é feita com um banco de filtros FIR. A qualidade do método desenvolvido é comprovada aplicando-o em mosaicos de texturas do tipo Brodatz e outros, obtendo-se excelentes índices de acerto quando comparado com inúmeros métodos de discriminação de texturas encontrados na literatura.

Abstract

In remote sensing images and other applications, such as medical imaging analysis, industrial inspection, etc. it is necessary, in many cases, the discrimination of textures. Intuitively, textures are related to tonal variations in the image, being a property of a region and not of a single point. These variations can be seen as local discontinuities and, in images, they are generally related to edges.

The wavelet transform provides a complete and stable representation of signals using edges in multi-scale. In this work, we develop an wavelet representation invariant by translations, with four directions, and, using this representation, we develop a method that aggregates and combines the energy responses of the wavelet transform in several scales and directions to obtain the discrimination of the textures.

This method has low computational cost, since it obtains its parameters from texture samples, and the discrete implementation of the wavelet decomposition is carried out with a FIR filter bank. The quality of the method developed is confirmed by applying it into mosaics of textures of the Brodatz type and others, obtaining excellent indexes of precision when compared to several other texture discrimination methods found in the literature.

Conteúdo

1	Introdução	1
1.1	Enquadramento Geral	1
1.2	Contribuição Original Desse Trabalho	4
1.3	Organização da Tese	4
1.3.1	Introdução	5
1.3.2	Metodologia	6
1.3.3	Aplicações	6
1.3.4	Perspectivas	6
1.3.5	Apêndices	7
2	Textura	9
2.1	Introdução	9
2.2	Definições de Textura	10
2.3	Aplicações da Análise de Texturas	13
2.4	Discriminação de Texturas	15
2.5	Sumário	19
3	Wavelets	21
3.1	Introdução	21
3.2	Análise Tempo Freqüência	22
3.3	As Origens	25
3.4	Transformada Wavelet Contínua em Uma Dimensão	26
3.5	Transformada Wavelet Diádica	27
3.6	Transformada Wavelet Discreta	28

3.6.1	Análise de Multiresolução	29
3.7	Sumário	30
4	Uma Família de Wavelets Diádicas	31
4.1	Introdução	31
4.2	Fundamentação	31
4.3	Wavelet Spline Diádica	34
4.4	Implementação no Caso Discreto	37
4.5	Algoritmo Discreto Para Reconstrução	39
4.6	Transformada Wavelet Diádica de Imagens	41
4.7	Condição de Admissibilidade Para $\psi^1, \psi^2, \psi^3, \psi^4$	43
4.8	Algoritmo Discreto	45
4.9	Sumário	48
5	Fundamentação e Modelo Experimental	49
5.1	Bordas em Multiescala	49
5.2	Multiresolução e Wavelets	51
5.3	Modelo Experimental	52
5.4	Etapas do Modelo Experimental	54
5.5	Algoritmo Combinatorial	61
5.7	Sumário	64
6	Aplicações do Modelo Experimental	65
6.1	Introdução	65
6.2	Trabalhos Anteriores	66
6.3	Segmentação de Pares de Texturas	68
6.3.1	Discussão dos Resultados	72
6.4	Segmentação de Mosaicos de Cinco Texturas e de Dez	74

Texturas	
6.4.1 Discussão dos Resultados	76
6.5 Uma Aplicação em Spring	81
6.6 Sumário	83
7 Conclusões e Trabalhos Futuros	85
7.1 Conclusões	85
7.2 Trabalhos Futuros	86
7.3 Considerações Finais	89
7.4 Sumário	89
Bibliografia	91
Apêndice A	101
Apêndice B	107

Capítulo 1

Introdução

1.1 Enquadramento Geral

Com o advento dos avanços tecnológicos possibilitados pelo aumento contínuo da capacidade de processamento dos computadores, imagens, que por muito tempo foram adquiridas exclusivamente por processos fotográficos, passaram mais e mais a serem objetos digitais, requerendo tecnologia específica ao seu tratamento e análise. Além disso, a rápida evolução das câmaras digitais, tanto no aspecto da qualidade como no custo, indica que, em breve, a fotografia digital terá qualidade equivalente à da fotografia convencional, por um preço acessível, tornando, em muitos casos, desnecessárias e obsoletas as imagens obtidas por processos químicos.

Imagens digitais são os objetos de estudo do Processamento Digital de Imagens, disciplina que tomou grande impulso a partir dos anos 60, aproveitando-se, principalmente, da evolução da informática e tornando-se um campo com inúmeras aplicações em diversas áreas. Seu objetivo principal é o de desenvolver ferramentas que possibilitem às máquinas não apenas automatizar procedimentos realizados pelas funções visuais dos seres humanos, mas também superar sua performance, pois são capazes de processar informação com muito mais rapidez e em maior quantidade. Contudo, como as máquinas não são capazes de pensar

ou tomar suas próprias decisões, há necessidade de desenvolver algoritmos que as auxiliem a executar as tarefas desejadas.

Dentre as várias características de uma imagem que são estudadas pelo Processamento Digital de Imagens, estão as **texturas**. O conceito de textura pode ser encontrado em diferentes contextos, definida de formas diversas e tendo inúmeras aplicações associadas. Essa diversidade torna impossível estabelecer uma única definição formal do conceito textura. De modo geral, regiões de texturas semelhantes estão relacionadas a características semelhantes e fornecem algum tipo de informação sobre o objeto representado pela imagem. Em certos casos, como por exemplo, em alguns tipos de imagens de radar, praticamente toda informação é textural.

Dentre as aplicações que se utilizam do conceito de textura, podemos destacar a análise de imagens médicas, o sensoriamento remoto, a inspeção industrial, o reconhecimento de formas de objetos tridimensionais a partir de imagens bidimensionais, a segmentação de documentos, entre outras.

A fim de caracterizar a complexidade das estruturas que podem estar presentes em uma imagem, tais como bordas e texturas, faz-se necessário buscar representações das imagens de modo que essas representações forneçam uma caracterização de atributos que, a princípio, não estariam claramente apresentados na representação original. Contudo, nem toda representação, por mais elegante que seja, matematicamente falando, vem a ser necessariamente uma boa representação. Isso, obviamente, dependerá do contexto e dos aspectos a serem realçados, ou seja, da finalidade e, portanto, da fidelidade buscada pela representação. De modo geral, uma representação adequada é um pré-requisito para se obter técnicas de processamento eficientes, tais como realce, filtragem e restauração, análise, reconstrução, etc.

Independentemente da natureza do problema que envolve o conceito textura, uma das melhores ferramentas para detecção e discriminação ainda é o olho humano. Desse modo, resultados obtidos em estudos psicofisiológicos da visão humana servem como referência ao desenvolvimento e construção de técnicas para a identificação e discriminação de texturas.

Esses estudos psicofisiológicos mostraram que o ser humano é capaz de diferenciar muitos tipos de texturas de modo pré-atentivo, isto é, diferenças entre texturas podem ser percebidas espontaneamente, sem a necessidade de se focar a atenção. Mostraram também que em outros tipos de textura isso não ocorre; são as chamadas texturas com estatísticas de segunda ordem [40]. Mas o mais interessante é que esses estudos indicaram, com fortes evidências, que a imagem na retina é decomposta pelo cérebro em vários canais de frequências espacialmente orientadas, e que existem mecanismos no córtex visual dos mamíferos que respondem a combinações de frequência e orientação dentro de um certo range [20]. Essa constatação levou inúmeros pesquisadores a desenvolverem modelos teóricos que mimetizam esses processos psicofisiológicos, acreditando que, devido à alta eficiência do sistema visual humano em perceber texturas, essa eficiência seria, de certo modo, herdada por esses modelos.

Uma das melhores ferramentas matemáticas que faz a intermediação entre representação espacial e frequencial é a Transformada de Fourier Janelada. Apesar de ter feito sucesso por um bom tempo, ela apresenta algumas limitações inerentes, quando analisando certos tipos de sinais e principalmente imagens. Sua resolução tempo-frequência é limitada pelo Princípio da Incerteza; baixas frequências terão valores altos se as janelas forem pequenas enquanto pulsos de curta duração terão localização muito pobre no tempo em janelas grandes.

As limitações da Transformada de Fourier Janelada motivaram a criação da Transformada Wavelet, a qual se caracteriza como uma decomposição em canais de multifrequência e tem se mostrado como uma poderosa ferramenta em inúmeras áreas de aplicação.

Dentro do contexto da transformada Wavelet, Mallat e Zhong [44] mostraram que a maior parte da informação sobre uma imagem concentra-se na relação entre as bordas que a compõe distribuídas em multiescalas. Eles desenvolveram uma decomposição (representação) em multiescalas, invariante por translação e, utilizando apenas as bordas¹ nas múltiplas escalas, mostraram

¹ (1) Bordas: Conceitualmente falando, são as regiões de transição entre os diferentes objetos em uma cena, caracterizadas em geral por mudanças acentuadas de brilho.

que é possível obter uma reconstrução estável do sinal com um mínimo de perdas, comprovando que as informações relevantes de um sinal são aquelas dadas por suas bordas.

As idéias concernentes à psicofisiologia da visão, juntamente com a decomposição por bordas em multiescalas apresentadas nos parágrafos acima, são utilizadas como os paradigmas que motivaram e fundamentaram a construção teórica e prática deste trabalho. Os conceitos envolvidos no processo teórico/prático necessário à sua concretização colocam este trabalho na confluência de várias áreas, tais como matemática pura, matemática aplicada, teoria de sinais, processamento de imagens e outras.

1.2 Contribuição Original Deste Trabalho

A premissa deste trabalho é que uma representação adequada da imagem é a condição primordial para se obter uma boa segmentação. A hipótese deste trabalho é a de que uma representação que considere escala, direcionalidade e a distribuição local da energia das bordas possibilitarão a obtenção de informações necessárias e suficientes para discriminar e separar inúmeras texturas.

Grosso modo, supõe-se que textura é uma combinação de escala, direcionalidade e densidade das componentes primitivas, sendo necessário quantificar/mensurar essas informações a fim de discriminar texturas. Essa quantificação, dentro da terminologia de reconhecimento de padrões é denominada extração de atributos.

Baseado nessa hipótese, desenvolve-se um modelo teórico matemático para uma decomposição Wavelet multiescala, invariante por translações e formada por quatro direções, juntamente com um modelo discreto, experimental, que agrega as respostas da transformada Wavelet nas várias escalas e direções com o objetivo de discriminar regiões de texturas homogêneas. A implementação desse modelo tem baixo custo computacional, pois todo treinamento e obtenção dos parâmetros são feitos em amostras tomadas das texturas, além disso, o

modelo matemático da família de Wavelets também foi construído de modo a ter baixo custo computacional em sua implementação discreta.

Os resultados obtidos pelo método desenvolvido, são comparados a vários outros métodos encontrados na literatura, mostrando índices de acerto maiores que os demais na maioria dos casos testados, confirmando a validade do modelo matemático teórico desenvolvido.

1.3 Organização da Tese

Este trabalho está dividido em quatro partes; uma parte introdutória, formada pelos três capítulos iniciais; uma parte metodológica, formada pelos dois capítulos subsequentes; uma parte de aplicações, formada pelo capítulo seis; e uma parte de perspectivas, apresentada no capítulo sete. Além disso, incluiu-se dois apêndices, um concernente à filtros não lineares, em particular os filtros de Volterra, e o outro contendo alguns programas e rotinas computacionais utilizadas na implementação prática das etapas descritas no modelo experimental. Pode-se dizer que a parte metodológica constitui o cerne deste trabalho. Uma visão geral dos capítulos é dada nessa seção.

1.3.1 Introdução

Capítulo 1: Introdução aos problemas, motivação deste trabalho, colocação dos paradigmas, contribuição original do trabalho.

Capítulo 2: Descrição do conceito de textura e suas possíveis definições. Mostra-se a necessidade e a atualidade de seu estudo e relatam-se diversas áreas onde o conceito de textura aparece. Apresentam-se várias abordagens citadas na literatura relacionadas à comparação de métodos de discriminação de texturas.

Capítulo 3: Apresenta-se uma introdução sobre a transformada Wavelet e relata-se um pouco sobre seu desenvolvimento histórico. Apresenta-se o conceito de Análise Tempo Freqüência e a forma como a transformada Wavelet se relaciona com esse conceito. Apresentam-se também outras abordagens que relacionam a Transformada Wavelet com esquemas de filtragem.

1.3.2 Metodologia

Capítulo 4: Desenvolve-se um modelo teórico matemático de uma família de Wavelets, invariantes por translação e com quatro direções, construídas tomando-se como base as funções Splines. Mostra-se sua relação com um banco de filtros discretos de resposta finita juntamente com a forma de implementá-las computacionalmente.

Capítulo 5: Apresenta-se a fundamentação teórica do trabalho. Mostra-se a relação de bordas em multiescalas como uma representação estável e completa de um sinal. Desenvolve-se um modelo experimental que implementa e otimiza as Wavelets construídas no capítulo quatro a fim de segmentar pares de texturas. Apresenta-se uma forma de generalizar esse método para mosaicos com mais de duas texturas e explicita-se detalhadamente como o modelo experimental opera.

1.3.3 Aplicações

Capítulo 6: Apresenta-se a aplicação do modelo experimental na segmentação² de mosaicos formados por pares de texturas e em mosaicos com cinco e dez texturas. Os resultados obtidos são comparados com inúmeros métodos propostos na literatura e dispostos em tabelas comparativas, demonstrando a boa performance do modelo aqui desenvolvido.

² Segmentar: separar os objetos de uma cena.

A fim de testar o desempenho do método como um gerador de atributos de texturas, utilizam-se as bandas de máxima separabilidade, obtidas com o modelo experimental, e faz-se uma classificação³ de um mosaico de cinco texturas utilizando o software Spring (“Sistema de Processamento de Informações Georeferenciadas”). O resultado encontra-se apresentado no final desse capítulo.

1.3.4 Perspectivas

Capítulo 7: Discute-se e sugere-se alternativas e possibilidades de trabalhos futuros utilizando o método desenvolvido e possíveis alterações dele. Apresenta-se uma abordagem não linear que, num certo sentido, engloba as tarefas executadas pelo modelo experimental aqui proposto. Essas não linearidades são conhecidas como filtros de Volterra. Apresentam-se também inúmeras possíveis abordagens que poderiam utilizar os filtros de Volterra para complementar o método aqui desenvolvido. Tais abordagens podem vir a ser utilizadas como idéias para trabalhos futuros.

1.3.5 Apêndices

No apêndice A, encontram-se descritas algumas propriedades de filtros não lineares, em especial sobre os filtros de Volterra. Sua inclusão deve-se ao fato de termos proposto várias possibilidades para trabalhos futuros, no capítulo 7, relacionadas à filtros de Volterra.

No apêndice B, apresentam-se as principais rotinas computacionais, utilizadas para implementar o modelo teórico para a segmentação dos mosaicos de texturas. Elas foram desenvolvidas utilizando o software IDL (Interactive Data Language) versão 5.2, da Research Systems Inc.

³ Classificar: rotular os objetos ou pontos de uma imagem.

Capítulo 2

Textura

2.1 Introdução

A natureza e os objetos ao nosso redor estão impregnados de texturas. Podemos encontrá-las em paredes, pisos, árvores, rochas, tecidos, na pele humana e dos animais, enfim, em praticamente toda matéria. Essa texturização, em muitos casos, é o que nos faz acreditar que um objeto é natural e não artificial, fazendo que, qualquer figura em computação gráfica, que queira parecer minimamente realista, tenha algum tipo de textura em sua superfície. Seu estudo já se estende por mais de três décadas; contudo, continua sendo tópico de enorme interesse. As pesquisas são motivadas em parte pela grande quantidade de texturas presentes em imagens e pela incrível capacidade do sistema visual humano em lidar com elas.

Até o momento, o instrumento mais eficiente para detectar e discriminar (diferenciar) texturas é o olho humano, servindo, na maioria das vezes como referência da exatidão dos algoritmos utilizados.

Os trabalhos pioneiros no assunto foram realizados por Julesz [37], [38], [39], [40] que estudou percepção de texturas extensivamente no contexto de discriminação de textura. A questão proposta por ele era: “Quando um par de texturas é discriminável, dado elas terem o mesmo brilho, contraste e cor?”, ou

seja, quais são os processos visuais que nos permitem destacar uma figura em relação ao conjunto, utilizando apenas suas características texturais?

A questão colocada por Julesz é a motivação básica entre os psicofisiologistas para estudarem percepção de textura. As respostas por eles encontradas são importantes, não só pelo fato já mencionado de que o sistema visual humano pode avaliar a performance dos algoritmos, mas pelo fato de inúmeros desses algoritmos serem desenvolvidos a partir dos resultados obtidos por esses estudos psicofisiológicos.

2.2 Definições de Textura

Devido à variedade de contextos e situações onde o termo textura aparece, não se tem uma definição formal desse conceito. Textura, em geral, está associada a padrões estruturais da superfície de objetos e, na maioria das vezes, é constituída por padrões repetitivos, cuja disposição pode ser periódica, quasi-periódica ou aleatória. Texturas naturais são, de modo geral, aleatórias, enquanto que texturas artificiais são freqüentemente determinísticas ou periódicas. Além disso, termos como fina, grossa, suave, áspera, granulada, alongada, regular e irregular estão associadas ao conceito de textura. Outra dificuldade que se coloca nessa questão é a da escala do objeto, pois, visto de perto, um objeto pode apresentar um tipo de textura, enquanto que, visto de longe, essa textura pode ser totalmente diferente. A figura 2.1 mostra alguns exemplos de texturas.

Tucierian e Jain [80] dão alguns exemplos de tentativas de definições, mostrando que as diferenças dependem da aplicação em particular em que o conceito de textura é utilizado. Segue abaixo alguns exemplos encontrados na literatura:

-“We may regard texture as what constitutes a macroscopic region. Its structure is simply attributed to the repetitive patterns in which elements or primitives are arranged according to a placement rule.” [79].

-“A region in an image has a constant texture if a set of local statistics or other local properties of the picture function are constants, slowly varying, or approximately periodic.” [75].

-“The image texture we consider is nonfigurative and cellular... An image texture is described by the number and types of its (tonal) primitives and the spatial organization or layout of its (tonal) primitives... A fundamental characteristic of texture: it cannot be analyzed without a frame of reference of tonal primitive being stated or implied. For any smooth gray-tone surface, there exists a scale such that when the surface is examined, it has no texture. Then as resolution increases, it takes on a fine texture and then a coarse texture.”[34]

-“Texture is defined for our purposes as attribute of a field having no components that appear enumerable. The phase relations between the components are thus not apparent. Nor should the field contain an obvious gradient. The intent of this definition is to direct attention of the observer to the global properties of the display-i.e. its overall “coarseness”, “bumpiness”, or “finess.” Physically, nonenumerable (aperiodic) patterns are generated by stochastic as opposed to deterministic processes. Perceptually, however, the set of all patterns without obvious enumerable components will include many deterministic (and even periodic) textures.” [73].

-“Texture is an apparently paradoxical notion. On the one hand, it is commonly used in the early processing of visual information, especially for practical classification purposes. On the other hand, no one has succeeded in producing a commonly accepted definition of texture. The resolution of this paradox, we feel, will depend on a richer, more developed model for early visual information processing, a central aspect of which will be representational systems at many different levels of abstraction. These levels will most probably include actual intensities at the bottom and will progress through edge and orientation descriptors to surface, and perhaps volumetric descriptors. Given these multilevel structures, it

seems clear that they should be included in the definition of, and in the computation of, texture descriptors.”[88].

Essa coleção de definições demonstra que a “definição” de textura tem sido formulada por diferentes pesquisadores, e pode-se verificar sua forte dependência em relação à aplicação particular a que se refere, ou seja, estão relacionadas a um determinado contexto e, em geral, acabam motivando o método que será utilizado para sua discriminação, não possuindo uma conformidade entre si.

2.3 Aplicações da Análise de Texturas

Inúmeras aplicações em diversas áreas necessitam utilizar-se de sistemas de reconhecimento de texturas. Abaixo, estão exemplificadas algumas.

Análise de Imagens Médicas

Imagens em medicina estão sendo obtidas por inúmeras técnicas não-intrusivas, tais como fotografias, raios x, ultra-som, tomografia, etc, e mesmo por técnicas intrusivas, tais como fotografias microscópicas ou biópsias. A análise dessas imagens tem um papel fundamental na obtenção de diagnósticos mais precisos. Tal análise depende de algumas propriedades, tais como cor, morfologia e textura. Em vários desses tipos de imagens, as propriedades texturais são determinantes para fins de diagnóstico. O câncer, por exemplo, é freqüentemente caracterizado por apresentar texturas características, podendo ser identificado utilizando-se diferentes técnicas de processamento de imagens médicas. A detecção precoce ainda é o modo mais efetivo de se reduzir a mortalidade, e a mamografia é a melhor forma de diagnóstico precoce. Técnicas relacionadas à textura têm sido aplicadas com sucesso na detecção de lesões em imagens mamográficas, na caracterização de lesões cervicais, entre outras [36], [54], [63].

Sensoriamento Remoto

Sensoriamento remoto é a utilização de sensores para a aquisição de informações sobre objetos ou fenômenos, sem que haja contato direto entre eles [50]. Técnicas de sensoriamento remoto incluem fotografias e medidas multi-espectrais feitas por satélite, inspeções feitas por sismógrafos ou sonares, entre outros.

Através do processo diferenciado de interação da radiação eletromagnética com a matéria nas diversas faixas do espectro dessa energia, obtém-se as imagens de sensoriamento remoto, cujo conteúdo fornece informações sobre alvos da superfície da Terra. Essas informações estão presentes, implicitamente, nas representações gráficas dessas imagens. O estudo dos atributos oriundos da resposta espectral dos alvos, juntamente com as técnicas de Processamento Digital de Imagens, indicam formas de discriminar e caracterizar os alvos, como por exemplo, tipos de culturas, áreas urbanas, reflorestamentos recentes ou antigos, desmatamento de florestas, queimadas, desertificação, etc, [55], [86], [25], [33].

Inspeção Industrial

Em processos industriais, a detecção de defeitos de produtos manufaturados é de importância crucial. Inspeção visual pode ser uma tarefa cansativa e com resultados ruins. Assim, uma forma de inspeção automática faz-se necessária. Em muitos casos, um critério de qualidade do material inspecionado pode ser a textura de sua superfície ou mesmo de sua micro-textura. Defeitos em objetos inspecionados costumam aparecer de modo muito claro no arranjo textural de sua superfície, possibilitando a detecção das bordas destes e sua conseqüente localização. Desse modo, análise de textura tem um papel importante em inspeção. Inúmeras abordagens para inspeção por texturas têm sido propostas [26], [70], [10].

Reconhecimento de Formas a Partir de Texturas

Pela análise de como os padrões de textura se comportam, a forma tridimensional de um objeto pode, em princípio, ser reconstruída a partir de imagens bidimensionais desse objeto [77], [87].

Segmentação de Documentos

Quando examinadas de uma distância suficiente, as regiões de um documento impresso são percebidas por seres humanos como textura. Para um idioma particular ou uma fonte específica, o texto pode se assemelhar a uma textura homogênea. Pelo uso de análise de texturas, pode ser possível segmentar regiões de texto e regiões de gráficos contidas em documentos escaneados. Tal segmentação é importante para se obter sucesso no reconhecimento ótico de caracteres e pode ser utilizada, por exemplo, na digitalização de documentos, jornais, artigos, etc [35] [70].

2.4 Discriminação de Texturas

A discriminação de texturas implica identificar regiões com texturas “uniformes” em uma dada imagem, a fim de destacá-las do contexto geral. Apesar de não haver uma definição formal, uma definição apropriada no contexto de segmentação poderia ser: “Uma região em uma imagem tem uma textura constante se um conjunto de estatísticas locais ou outras propriedades locais da figura são constantes, ligeiramente variantes ou aproximadamente periódicas” [75]. A diversidade de texturas naturais e artificiais torna impossível uma definição universal e também a existência de uma única abordagem que englobe todos os casos. Desse modo, os métodos para explorar texturas são diversos e refletem os diferentes conceitos de textura que aparecem em aplicações de diversas áreas. Abaixo, estão descritos alguns métodos importantes:

Métodos Estatísticos

Uma das qualidades que definem uma textura é a distribuição espacial de seus tons de cinza. Métodos estatísticos são utilizados na obtenção de informações sobre texturas explorando as estatísticas dos níveis de cinza dos pixels. Estas podem ser estatísticas de primeira ordem, como por exemplo, histograma, média e variância, ou as de ordens altas, principalmente de segunda ordem, que capturam correlações locais dos níveis de cinza dos pixels da imagem. Os métodos mais comumente usados são baseados na matriz de co-ocorrência dos níveis de cinza [33] [34], a partir da qual atributos de textura são derivados. Provavelmente, esse é o método mais popular e referenciado para discriminação de texturas. Ele foi introduzido por Haralick *et al* [33], em 1973. Nesse método, as frequências relativas de pixels com mesmo tom de cinza e que ocupam determinadas posições são computadas e armazenadas na matriz de co-ocorrência P . Se houver M tons de cinza na imagem, P será de tamanho $M \times M$. Se M é grande, o número de pares de pixels contribuindo para cada elemento p_{ij} em P será pequeno. Se M é muito pequeno, então a maior parte da informação sobre a textura terá sido perdida na quantização da imagem e P não dará informação adequada. Ohanian e Dudes [58] reportaram que $M = 8$ era uma escolha apropriada para sub-imagens de tamanho 32×32 , contudo, outros autores usam outras quantizações (em Dutra e Huber [25], $M=16$ e sub-imagens 11×11). Além disso, usualmente consideram-se pares de pixels em orientações de 0, 45, 90 e 135 graus.

Haralick [34] sugere o uso de 14 atributos extraídos de P ; contudo, com o passar do tempo, verificou-se que quatro desses atributos são os mais interessantes, [58], [76]. São estes: Segundo Momento Angular (SMA), Contraste (COM), Correlação (COR) e Entropia (ENT). Eles são dados por:

$$SMA = \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} (p_{ij})^2$$

$$COM = \sum_{n=0}^{M-1} N^2 \left\{ \sum_{|i-j|=n} p_{ij} \right\}$$

$$\text{COR} = \frac{\sum_{i=0}^{M-1} \sum_{j=0}^{M-1} ijP(i, j) - \mu_x \mu_y}{\sigma_x \sigma_y}$$

$$\text{ENT} = - \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} P(i, j) \log P(i, j)$$

onde

$$\mu_x = \sum_{i=0}^{M-1} iP(i, *),$$

$$\mu_y = \sum_{j=0}^{M-1} jP(*, j),$$

$$\sigma_x = \sqrt{\sum_{i=0}^{M-1} (i - \mu_x)^2 P(i, *)},$$

$$\sigma_y = \sqrt{\sum_{j=0}^{M-1} (j - \mu_y)^2 P(*, j)},$$

$$P(i, *) = \sum_{j=0}^{M-1} P(i, j),$$

$$P(*, j) = \sum_{i=0}^{M-1} P(i, j).$$

e

P é a matriz de co-ocorrência.

Da mesma forma que a quantização, a escolha dos atributos dependem do tipo de aplicação e dos autores; por exemplo, em Dutra e Huber [25], são utilizados os atributos: Energia, Entropia, Máxima Probabilidade, Contraste, Correlação e Homogeneidade.

De modo geral, o uso dos quatorze atributos propostos por Haralick se torna inviável, pois demandam um enorme esforço computacional, tanto no sentido do processamento numérico quanto no do armazenamento das informações obtidas e os resultados ficam aquém do esperado.

Outros métodos estatísticos podem, por exemplo, ser baseados no cálculo da função de autocorrelação. Ela avalia a regularidade no sentido da textura ser fina ou grossa. A função de autocorrelação, normalizada, de uma imagem I

calculada no ponto (x, y) em uma região de tamanho NxN pode ser estimada como [49]:

$$\rho(x, y) = \frac{\sum_{u=0}^N \sum_{v=0}^N I(u, v) I(u+x, v+y)}{\sum_{u=0}^N \sum_{v=0}^N I^2(u, v)}.$$

Esta função está relacionada ao tamanho das primitivas componentes das texturas, isto é, se fina ou grossa. Além disso, também se relaciona ao espectro de potência da Transformada de Fourier da imagem [24].

Métodos Estruturais

Descrevem texturas como sendo compostas de primitivas (texels) bem definidas, que são arranjadas conforme alguma regra sintática [27]. Como este modelo permite descrever apenas texturas regulares, as regras são freqüentemente estendidas para tornarem-se estatísticas, o que oferece uma liberdade maior na descrição.

Métodos Baseados em Modelos

Constroem um modelo gerador ou estocástico de texturas. Os parâmetros do modelo são estimados na imagem e atuam como atributos. Os que obtiveram mais sucesso em muitas aplicações são os “*random field models*” (modelos auto-regressivos, *Markov random fields*, *Gibbs-Markov models*) [8] [9]. Alguns outros modelos como, por exemplo, modelos fractais, se inserem nessa categoria.

Métodos de Transformadas

Representam uma imagem em uma nova forma, na qual as características da textura se tornam mais facilmente acessíveis. Exemplos disso são os métodos espectrais baseados na transformada de Fourier. Uma subclasse importante é formada pelos métodos de multiresolução, que transformam imagens em uma representação que separa atributos em diferentes escalas de resolução tais como a Transformada de Gabor e a Transformada Wavelet.

Métodos de Multirresolução

São importantes devido à sua relação com o sistema visual humano. Estudos psicofisiológicos apresentam evidências de que humanos processam imagens em um modo multiescala. Os “*early processing*” no cérebro realizam um tipo de análise espaço-freqüência. O córtex visual possui células distintas que respondem a diferentes freqüências e orientações [5]. Esse processamento em multiescala que os seres humanos aplicam com muito sucesso, a fim de reconhecer e identificar texturas, é uma motivação para se proceder a uma análise de texturas em multiescala. Exemplos desses são as decomposições espaço-escala da Transformada de Gabor e a Transformada Wavelets [85] [26] [35] [81]. Esta tese se insere nessa classe.

2.5 Sumário

Nesse capítulo, fez-se uma descrição geral sobre o conceito de textura e suas possíveis definições. Apresentou-se várias aplicações que se utilizam da análise de textura e o papel que o conceito de textura exerce em cada uma delas, juntamente com abordagens relacionadas à discriminação de texturas.

De modo geral os métodos aqui descritos pertencem à classe dos métodos que utilizam bancos de filtros que extraem os atributos relacionados com textura e com base nestes atributos, separam-se os objetos da cena.

Em outra classe de algoritmos, primeiro separam-se os objetos para posteriormente rotulá-los com base na textura. A medição das propriedades das texturas é, no entanto, semelhantes aos já descritos.

Capítulo 3

Wavelets

3.1 Introdução

Por muito tempo, a transformada de Fourier foi a ferramenta soberana em processamento de sinais, deixando pouco espaço para novas incursões matemáticas nesse ramo. Até os anos 70, sinais estavam relacionados principalmente à voz e outros sons, que, de modo geral, eram modelados como realizações de processos gaussianos. Como resultado, algoritmos lineares eram considerados ótimos sobre todos os outros procedimentos, e, com a hipótese de estacionariedade, ficava-se restrito à classe dos operadores de convolução que são diagonalizados pela transformada de Fourier.

Nos anos 80, essa situação se alterou completamente, com o desenvolvimento do processamento de imagens. Processos gaussianos modelam imagens de modo muito pobre, visto que estruturas transientes, tais como bordas, são freqüentemente mais importantes que propriedades estacionárias. Desse modo, algoritmos não-lineares se tornaram inevitáveis, abrindo o processamento de sinais a abordagens matemáticas mais sofisticadas.

Uma abordagem que aflorou nesse período foi a construção de representações esparsas de sinais (funções), processos e operadores. Esse tipo de representação está na raiz de muitos problemas de processamento de sinais e informação. Uma representação esparsa caracteriza-se como uma aproximação com poucos parâmetros, podendo ser obtida a partir de uma expansão em uma

base de funções ou, de modo mais geral, num “dicionário” de funções [48]. Sob tais representações, processos não-lineares de grande complexidade podem freqüentemente ser reduzidos a algoritmos simples e rápidos. Meyer [52], associou tais representações com Análise Harmônica, enquanto que Daubechies [19] e Donoho [21] mostraram o impacto das bases de wavelets em análise numérica e estatística.

A transformada de Fourier expande um sinal usando um conjunto de funções periódicas. Os coeficientes resultantes são, em geral, realizados como uma combinação de dois sinais; a magnitude do sinal, que contém o conteúdo de freqüências, pois as funções representam diferentes freqüências; e a fase do sinal, que contém a informação espacial, mas de um modo implícito que não pode ser interpretado diretamente; além disso, a transformada de Fourier assume que o sinal a ser analisado seja de duração infinita ou pelo menos periódico. Obviamente, essa condição nem sempre é satisfeita quando se trabalha com sinais do mundo real. Considere, por exemplo, um trecho de uma música. Ele consiste de várias notas com freqüências particulares que são tocadas durante um tempo limitado. Esse exemplo aponta para a necessidade de uma representação tempo-freqüência caso o sinal a ser analisado contenha freqüências que existem apenas durante um intervalo de tempo limitado. Veja o exemplo dado na figura 3.1.

3.2 Análise Tempo Freqüência

Uma maneira de obter uma representação, tal qual apontada pelo parágrafo anterior, é colocar uma janela, localizada sobre o sinal, que seja deslocada ao longo do eixo do tempo e executar a transformada de Fourier dentro dessa janela. Isto dará origem à Transformada de Fourier Janela (WFT):

$$Sf(u, \xi) = \int f(t) g(t-u) e^{-i\xi t} dt$$

onde g é a janela no domínio do tempo. Tomando,

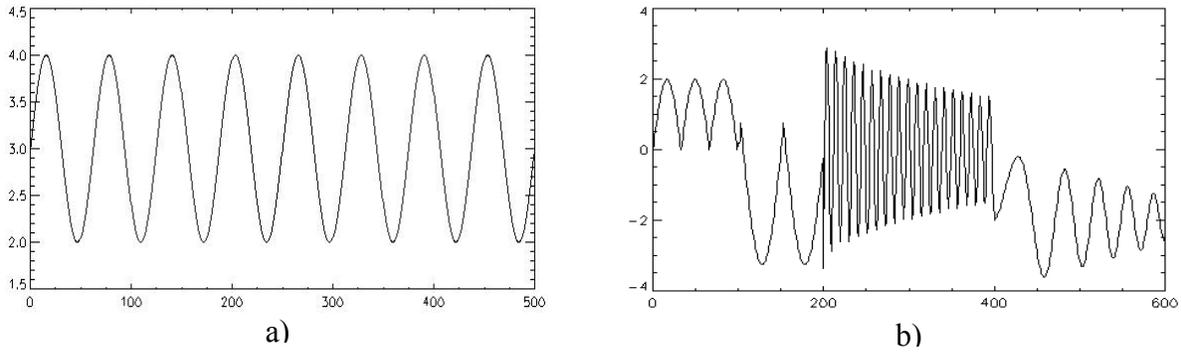


Figura 3.1: Dois sinais com diferentes características tempo-freqüência. a) O sinal é periódico, e a transformada de Fourier é capaz de analisar seu conteúdo de freqüências. b) A análise de Fourier não é apropriada, pois o sinal exibe diferentes comportamentos na freqüência em diferentes janelas de tempo.

$$g_{u,\xi}(t) = g(t-u)e^{i\xi t}$$

tem-se que a energia de $g_{u,\xi}(t)$ está concentrada na vizinhança de u , sobre um intervalo de tamanho σ_t , medido pelo desvio padrão de $|g|^2$, enquanto que tomando a transformada de Fourier \hat{g} de g ,

$$\hat{g}_{u,\xi}(\omega) = \hat{g}(\omega - \xi)e^{-iu(\omega - \xi)}$$

verifica-se que a energia de $\hat{g}_{u,\xi}$ está localizada ao redor da freqüência ξ , sobre um intervalo de tamanho σ_ω , que mede a região onde $\hat{g}(\omega)$ possui valores não desprezíveis. No plano tempo-freqüência (t, ω) , conforme figura 3.2a, o espalhamento da energia de $g_{u,\xi}(t)$ é representada pelos retângulos. Esses retângulos são centrados em (u, ξ) e possuem área igual a $\sigma_t \sigma_\omega$. Apesar de a forma desses retângulos poder variar, dependendo do tamanho da janela g , no tempo, o princípio da incerteza estabelece que sua área satisfaz,

$$\sigma_t \sigma_\omega \geq \frac{1}{2}.$$

Esta área é mínima quando g é uma gaussiana, e, nesse caso, os átomos $g_{u,\xi}$ são chamados de funções de Gabor.

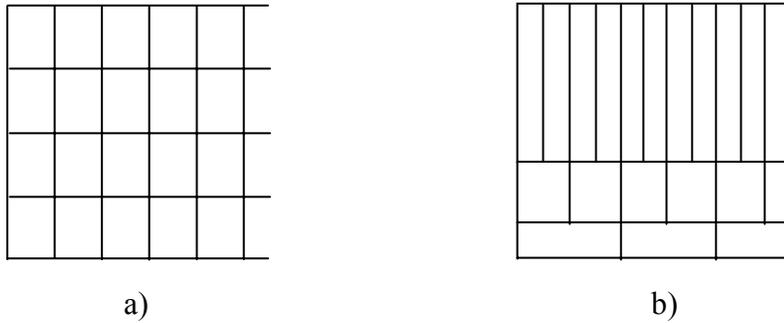


Figura 3.2: “Tiling” do plano tempo freqüência. a) Como obtido pela WTF. b) Como obtido pela transformada Wavelet.

É desejável, contudo, estudar propriedades de um sinal que variam lentamente (baixas freqüências) por um longo período de tempo, e estudar propriedades que variam muito rapidamente (altas freqüências) por curtos períodos de tempo. Uma forma esquemática desse tipo de estudo é dada pela relação tempo freqüência, mostrada na figura 3.2b e uma forma de implementá-lo é pela transformada Wavelet.

A transformada Wavelet expande uma função usando um conjunto de funções wavelets. Os coeficientes resultantes contêm o conteúdo de freqüências, pois as funções wavelets representam diferentes escalas (ou freqüências). As funções wavelets são também localizadas espacialmente, o que significa que a expansão é local, fazendo que os coeficientes sejam computados para todos os pontos do sinal original e, portanto, a informação espacial continua intacta.

Devido à enorme variedade de áreas e aplicações associadas ao conceito de wavelets, Sweldens [78] procurou definir algumas características, isto é, a idéia do que são wavelets e porque são desejáveis. São três as características:

- As wavelets são blocos construtores de funções em geral, isto é, uma função pode ser representada como uma série infinita de funções wavelets. Se $\Psi = \{\psi_\lambda : \lambda \in \Lambda\}$ é uma base de wavelets, f é uma função e $c = \{c_\lambda : \lambda \in \Lambda\}$, então,

$$f = \sum_{\lambda} c_{\lambda} \psi_{\lambda}$$

- As wavelets têm localização espaço-freqüencial. Por localização espacial entende-se que a maioria da energia da wavelet está restrita a uma região finita. Idealmente, a wavelet teria suporte compacto. Por localização freqüencial entende-se que a transformada de Fourier da wavelet é localizada, isto é, a maioria das freqüências está localizada em uma determinada banda.
- As wavelets têm algoritmos rápidos. Deve ser possível e fácil implementar funções wavelets num computador. Tipicamente, o algoritmo deve ter um custo computacional linear ou linear-logarítmico.

3.3 As Origens

O nome “Wavelet” foi introduzido, nos anos oitenta, por Morlet [53], um geofísico francês. Os dados sísmicos que ele estudava exibiam mudanças bruscas no conteúdo das freqüências, comportamento esse para o qual a análise de Fourier não era uma ferramenta suficientemente adequada. Assim, ele passou a investigar a existência de funções que tivessem suporte compacto tanto no domínio do tempo como da freqüência, chamando tais funções de wavelets. Em seguida, Grossman forneceu as bases matemáticas para as idéias de Morlet [30], dando origem ao ferramental matemático conhecido como teoria de wavelets.

Voltando às origens verifica-se que essa não é de todo uma teoria nova. Ao invés disso, ela é o cruzamento e a generalização de conceitos conhecidos em várias áreas, tais como, geofísica, análise e compressão de sinais (e imagens), física e matemática (estatística). O passo inicial foi a transição da análise em freqüências de Fourier (1807) para a análise em escala, dado pelo matemático alemão Alfred Haar (1910) [31]. É onde se observa a primeira relação com wavelets. Sua wavelet tinha suporte compacto, mas não era contínua.

Alguns resultados essenciais no desenvolvimento da análise wavelet foram obtidos pela cooperação de Ingrid Daubechies, uma física teórica belga e Stephane Mallat, um francês com formação em matemática e engenharia elétrica. Daubechies construiu famílias de wavelets ortonormais com suporte compacto [18], enquanto que Mallat [47] relacionou a transformada Wavelet com o conceito de decomposição de um sinal em multirresolução, fornecendo uma forma de implementar a transformada Wavelet de modo rápido e com filtros discretos. O conceito de multirresolução já era bem conhecido na área de compressão de imagens. Essas pesquisas resultaram num ferramental para análise de sinais com aplicações em física, reconhecimento de voz, codificação de imagens, reconhecimento e segmentação de imagens, *denoising*, estimação de densidade e outras.

Para dar uma visão global do que se tornou a teoria de wavelets, seria necessário mais que um único livro. Alguns bons textos sobre o assunto são, Mallat [46], Daubechies [17], Chui [12] [13], Vetterli [83], entre outros. Nesse capítulo, iremos nos limitar a introduzir os conceitos que serão necessários ou que de algum modo se relacionam com a teoria de wavelets necessária aos capítulos posteriores.

3.4 Transformada Wavelet Contínua em Uma Dimensão

Seja $\psi_s(x) = \frac{1}{s} \psi\left(\frac{x}{s}\right)$. A transformada Wavelet (WT) de uma função $f(x) \in L^2(\mathbb{R})$ na escala s e posição x é dada pela convolução,

$$W_s f(x) = f * \psi_s(x). \quad (3.1)$$

A função ψ é chamada de wavelet mãe e a equação (3.1) corresponde à wavelet mãe dilatada pelo fator s e transladada pelo fator correspondente da convolução. Grossman e Morlet [30] mostraram que se a transformada de Fourier

da wavelet $\psi(x)$ satisfaz $\hat{\psi}(0) = 0$, então a transformada Wavelet satisfaz uma equação de conservação de energia, e $f(x)$ pode ser reconstruída a partir de sua WT. Quando a escala s cresce, o suporte de $\psi_s(x)$ decresce, de modo que a transformada Wavelet $W_s f(x)$ captura detalhes finos; e quando a escala s decresce, o suporte de $\psi_s(x)$ cresce, e a $W_s f(x)$ captura detalhes em escalas maiores. A wavelet $\psi(x)$ pode ser interpretada como a resposta a impulso de um filtro passa banda, e a transformada Wavelet $W_s f(x)$ como uma convolução com um filtro passa banda que é dilatado.

3.5 Transformada Wavelet Diádica

Em reconhecimento de padrões, é importante construir representações de sinais que sejam invariantes por translações, pois, quando um padrão é transladado, seus descritores numéricos devem também ser transladados, mas não modificados. Um padrão cuja representação depende da sua localização, torna-se difícil de ser encontrado e perde sua relação com estruturas locais. A transformada Wavelet e a transformada de Fourier janelada contínuas fornecem representações invariantes por translação; entretanto, a amostragem uniforme do parâmetro de translação destrói sua invariância por translação [46], o que a torna inconveniente para um esquema padrão de extração de atributos. A transformada Wavelet diádica apresentada a seguir é uma representação invariante por translação. Essa representação de um sinal é altamente redundante, isto é, as bandas são altamente correlacionadas; contudo, cabe salientar que, cada uma das saídas geradas pela decomposição wavelet, traz alguma informação adicional, muitas vezes não presente em outros canais, e é esse fato que será explorado pelo modelo experimental desenvolvido no capítulo 5.

Para construir uma representação que seja invariante por translação, na definição da transformada Wavelet diádica, a escala é discretizada, enquanto que o parâmetro de translação x não é. A escala é discretizada ao longo da seqüência

diádica $(2^j)_{j \in \mathbb{Z}}$, e a WT de $f(x)$ na escala 2^j e na posição x é dada pela convolução

$$W_{2^j} f(x) = f * \psi_{2^j}$$

A transformada de Fourier de $W_{2^j} f(x)$ é dada por

$$\hat{W}_{2^j} f(w) = \hat{f}(w) \hat{\psi}(2^j w)$$

Se existirem constantes $A_1, B_1 > 0$ tais que

$$\forall w \in \mathbb{R} - \{0\}, A_1 \leq \sum_{j=-\infty}^{+\infty} \left| \hat{\psi}(2^j w) \right|^2 \leq B_1 \quad (3.2)$$

então fica assegurado que todo o eixo de frequências é coberto por dilatações de $\hat{\psi}(w)$ por $(2^j)_{j \in \mathbb{Z}}$, de modo que $\hat{f}(w)$ e $f(x)$ podem ser recuperadas a partir de sua WT diádica [42]. A expressão (3.2) é chamada de Condição de Admissibilidade. A wavelet de reconstrução $\chi(x)$ é qualquer função cuja transformada de Fourier satisfaça

$$\sum_{j=-\infty}^{+\infty} \hat{\psi}(2^j w) \hat{\chi}(2^j w) = 1.$$

Desse modo, a função $f(x)$ pode ser recuperada a partir de sua WT diádica, [46], com a soma

$$f(x) = \sum_{j=-\infty}^{+\infty} W_{2^j} f * \chi_{2^j}(x).$$

3.6 Transformada Wavelet Discreta

A fim de obter a transformada Wavelet de sinais discretos, foram desenvolvidas inúmeras abordagens. De modo geral, utilizam-se parâmetros de dilatação e translação que não fazem uso da transformada Wavelet sobre toda a faixa de valores, mas sim sobre valores discretos.

Essencialmente, a transformada Wavelet é uma convolução de uma aplicação $f(t)$ com uma wavelet escalonada $\psi_a(x)$ para todos os valores de a . Pelo

teorema da convolução, isto se torna uma multiplicação no domínio de frequência. Logo, para sinais discretos, a transformada Wavelet pode ser obtida calculando-se a transformada rápida de Fourier (FFT), fazendo-se uma multiplicação e calculando-se a transformada rápida inversa de Fourier (IFFT). A complexidade computacional desse procedimento é muito alta, principalmente para imagens. É, portanto, desejável discretizar a transformada, restringindo-se os parâmetros de dilatação a e de translação b a uma grade discreta, isto é, $a = a_0^m$ e $b = n b_0 a_0^m$, com $m, n \in \mathbb{Z}$ e $a_0 > 1$, $b_0 > 0$. A escolha mais popular tem sido $a_0 = 2$ e $b_0 = 1$.

Uma forma de facilitar a discretização e obter algoritmos rápidos torna-se possível ao se fazer uma escolha adequada da wavelet mãe. Se algumas relações de escalas fizerem parte das propriedades da wavelet mãe, então a discretização se torna uma filtragem de fácil implementação e com complexidade computacional baixa. Essas relações de escala estão no cerne de outras abordagens utilizadas para a construção de wavelets.

3.6.1 Análise de Multirresolução

A análise de multirresolução consiste de uma seqüência de sucessivos espaços de aproximação V_j . Formalmente, uma seqüência $\{V_j\}_{j \in \mathbb{Z}}$ de subespaços fechados de $L^2(\mathbb{R})$ é uma análise de multirresolução se as seguintes propriedades são satisfeitas [46]:

$$\text{i -) } \forall (j, k) \in \mathbb{Z}^2, f(t) \in V_j \Leftrightarrow f(t - 2^{-j}k) \in V_j,$$

$$\text{ii -) } \forall j \in \mathbb{Z}, V_j \subset V_{j+1},$$

$$\text{iii -) } \forall j \in \mathbb{Z}, f(t) \in V_j \Leftrightarrow f(2t) \in V_{j+1},$$

$$\text{iv -) } \lim_{j \rightarrow -\infty} V_j = \bigcap_{j=-\infty}^{+\infty} V_j = \{0\}$$

$$\text{v -) } \lim_{j \rightarrow +\infty} V_j = \text{fecho} \left(\bigcup_{j=-\infty}^{+\infty} V_j \right) = L^2(\mathbb{R})$$

vi -) Existe θ tal que $\{\theta(t-n)\}_{n \in \mathbb{Z}}$ é uma base de Riesz de V_0 .

Intuitivamente, a propriedade (i) significa que V_j é invariante por qualquer translação proporcional à escala 2^j . A propriedade (ii) implica que uma aproximação numa resolução 2^j contém toda informação necessária para computar uma aproximação na resolução 2^{j+1} .

A análise de multirresolução fornece um método para a construção de bases ortonormais de wavelets, além de mostrar como entender tais bases [47]. Sua popularidade se deve principalmente ao fato que sua implementação discreta é fácil e com algoritmos rápidos.

Dentro do contexto da Análise de Multirresolução, as wavelets mais populares são, provavelmente, a wavelet de Mallat [47] e a família de wavelets desenvolvida por Daubechies [18]. Suas wavelets são ortonormais e possuem suporte compacto, caracterizando-se por um conjunto de filtros com tamanho a partir de quatro coeficientes. Podemos encontrá-las implementadas em inúmeros softwares de uso comercial nos dias de hoje, além disso, essas wavelets têm um caráter didático bastante interessante. Sua deficiência, no contexto de extração de atributos, é devido ao fato delas não serem uma representação invariante por translação e sua implementação discreta ser na forma de um algoritmo piramidal.

3.7 Sumário

Nesse capítulo, apresentou-se uma breve introdução às idéias e conceitos básicos da teoria de wavelets, juntamente com um pouco de sua história.

Apresentou-se também o conceito de análise tempo-freqüência e como a transformada Wavelet se encaixa nesse conceito. Definiu-se a transformada Wavelet contínua, as transformadas Wavelets diádicas contínuas e discretas, além do conceito de multirresolução e as wavelets de Daubechies, justificando-se o motivo de não usá-las essas e sim as diádicas.

Capítulo 4

Uma Família de Wavelets Diádicas

4.1 Introdução

Nesse capítulo, desenvolve-se uma família de wavelets a partir das splines, a fim de poder computar a transformada wavelet diádica discreta por um banco de filtros, com um algoritmo rápido. Para isso, deve-se construir a wavelet de modo apropriado.

O ferramental matemático utilizado na construção da família de wavelets é baseado nos trabalhos de Mallat e Zhong [44], [45]. Contudo, várias modificações foram feitas a fim de estender esse ferramental. Além da invariância por translação e das multiescalas, busca-se tornar a decomposição mais apropriada para descrever o comportamento das texturas. Com esse objetivo, incluem-se duas direções além das tradicionais; horizontal e vertical. Mostra-se também como essa família de wavelets relaciona-se com um banco de filtros e como é feita sua implementação discreta.

4.2 Fundamentação

Conforme Mallat e Zhong [44], pode-se tomar a wavelet $\psi(x)$ como sendo a derivada de primeira ordem ou de segunda ordem de uma função $\theta(x)$, isto é,

$$\psi^{(1)}(x) = \frac{d\theta(x)}{dx} \quad (4.1)$$

ou

$$\psi^{(2)}(x) = \frac{d^2\theta(x)}{dx^2} \quad (4.2)$$

onde $\theta(x)$ é a resposta a impulso de um filtro passa-baixa. $\theta(x)$ é uma função cuja transformada de Fourier tem a energia concentrada em baixas frequências. Uma tal função é chamada de função suavizante. Um exemplo desse tipo de função é a gaussiana,

$$\theta(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(\frac{-x^2}{2\sigma^2}\right)$$

Escolhendo as wavelets como em (4.1) ou (4.2), segue que

$$W_s^1 f(x) = f * \left(s \frac{d\theta_s}{dx} \right)(x) = s \frac{d}{dx} (f * \theta_s)(x) \quad (4.3)$$

$$W_s^2 f(x) = f * \left(s^2 \frac{d^2\theta_s}{dx^2} \right)(x) = s^2 \frac{d^2}{dx^2} (f * \theta_s)(x) \quad (4.4)$$

onde $\theta_s(x) = \frac{1}{s} \theta\left(\frac{x}{s}\right)$. As equações (4.3) e (4.4) são respectivamente a primeira e a segunda derivadas do sinal suavizado na escala s . Assim, a transformada wavelet $W_s^1 f(x)$ é proporcional à derivada primeira de $f(x)$ suavizada por $\theta_s(x)$, e, portanto, os extremos locais de (4.3) correspondem aos cruzamentos por zeros (*zero-crossings*) de (4.4). A detecção dos extremos locais da transformada wavelet em (4.3) é essencialmente equivalente a um detector de bordas de Canny [7], enquanto que a detecção dos cruzamentos por zeros de (4.4) corresponde a um detector de bordas de Marr-Hildreth [50] (em ambos os casos, $\theta(x)$ é tomada como uma gaussiana).

Mallat e Zhong [45] mostraram que a detecção dos máximos em (4.3) é superior a detecção dos cruzamentos por zeros em (4.4), (veja a figura 4.1) por dois motivos,

1- um ponto de inflexão de $f * \theta_s(x)$ pode ser um máximo ou um mínimo do módulo de sua derivada primeira. Os máximos desse módulo correspondem a variações abruptas de $f * \theta_s(x)$, enquanto que os mínimos correspondem a variações suaves. Esses dois tipos de pontos de inflexão podem ser distinguidos um do outro verificando se um extremo de $|W_s^1 f(x)|$ é um máximo ou um mínimo, mas eles não podem ser diferenciados pelos cruzamentos por zeros de $|W_s^2 f(x)|$.
Veja figura 4.1.

2- cruzamentos por zeros não fornecem uma representação estável do sinal, pois pequenas perturbações de $f(x)$ em uma região, onde $|W_s^2 f(x)|$ tem uma pequena amplitude, pode modificar consideravelmente as posições dos cruzamentos por zeros.

A fim de poder computar a transformada wavelet diádica discreta por um banco de filtros do tipo FIR, deve-se construir a wavelet de modo apropriado. Para isso, sejam h e g um par de filtros que possuem resposta de impulso finita (FIR). Suponha que h seja um filtro passa baixa cuja função de transferência, $H(\omega)$ satisfaz $\hat{H}(0) = 1$ e $G(\omega)$ a função de transferência de g . Como nos casos de bases wavelets ortogonais e biortogonais [46], constrói-se uma função de escala cuja transformada de Fourier é dada por,

$$\hat{\phi}(\omega) = \prod_{p=1}^{+\infty} \hat{H}(2^{-p} \omega) = \hat{H}\left(\frac{\omega}{2}\right) \hat{\phi}\left(\frac{\omega}{2}\right) \quad (4.5)$$

Supondo que $\phi \in L^2(\mathbb{R})$, a transformada de Fourier da wavelet correspondente ψ será dada por

$$\hat{\psi}(\omega) = \hat{G}\left(\frac{\omega}{2}\right) \hat{\phi}\left(\frac{\omega}{2}\right) \quad (4.6)$$

e, nesse caso, segue que ϕ e ψ possuem suporte compacto, pois h e g tem um número finito de coeficientes não nulos.

Essa construção inicial, juntamente com as B-splines, serão utilizadas no parágrafo seguinte para a construção de uma família de wavelets com propriedades adequadas e a determinação dos filtros discretos h e g ($H(\omega), G(\omega)$).

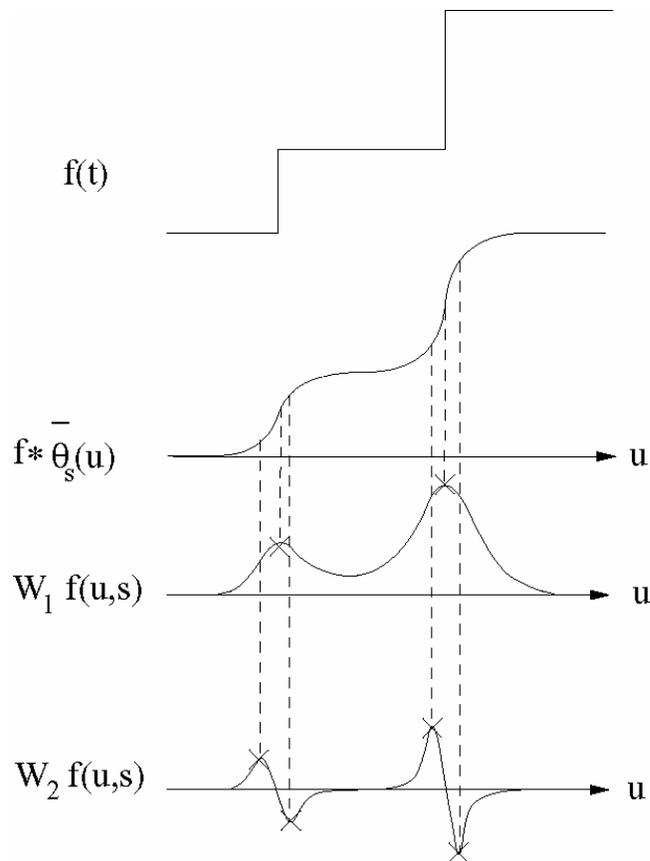


Figura 4.1: Se $\psi = -\theta'$ então $W_1 f(u,s) = s \frac{d}{du} (f * \bar{\theta}_s)(u)$ tem módulo máximo em pontos de variação abrupta de $f * \bar{\theta}_s(u)$. Se $\psi = \theta''$, então os módulos máximos de $W_2 f(u,s) = s^2 \frac{d^2}{du^2} (f * \bar{\theta}_s)(u)$ correspondem a máximos locais das curvas.

4.3 Wavelet Spline Diádica

B-splines é uma família de funções cujas propriedades possibilitam estabelecer relações conforme (4.5), além disso, possuem uma relação intrínseca com as Gaussianas. Esses fatos serão utilizados nas construções a seguir.

Uma B-spline de ordem n é definida como sendo a convolução de n splines de ordem zero,

$$B^n(x) = B^0 * B^{n-1}(x) = B^0 * B^0 * \dots * B^0 \quad (n+1 \text{ vezes})$$

onde $B^0(x)$ é a função característica do intervalo $(0,1]$ (*box function*). O suporte da B-spline de ordem n é o intervalo $[0, n]$ e sua transformada de Fourier é dada por

$$\hat{B}^n(\omega) = e^{-ik\frac{\omega}{2}} \left(\frac{\text{sen} \frac{\omega}{2}}{\frac{\omega}{2}} \right)^{n+1} \quad \text{onde } k = \begin{cases} 1 & \text{se } n \text{ é par} \\ 0 & \text{se } n \text{ é ímpar} \end{cases}$$

Se ϕ é uma B-spline, então a relação de escala (4.5) diz que

$$\hat{H}(\omega) = \frac{\hat{B}^n(2\omega)}{\hat{B}^n(\omega)} = e^{-\frac{ik\omega}{2}} \left[\cos \frac{\omega}{2} \right]^{n+1}$$

Usando a fórmula de Euler e expansão binomial é possível calcular os coeficientes FIR de $\hat{H}(\omega)$. Eles são dados por

$$h_k = \begin{cases} \frac{1}{2^{n+1}} \binom{n+1}{\frac{n+1}{2} + k} & ; \quad |k| \leq n+1 \\ 0 & \text{caso contrário} \end{cases} \quad \text{para } n \text{ ímpar}$$

onde $\binom{n}{k} = \frac{n!}{k!(n-k)!}$ são os coeficientes binomiais e

$$h_k = \begin{cases} \frac{1}{2^{n+1}} \binom{n+1}{\frac{n}{2} + 1 + k} & ; \quad -\frac{n}{2} - 1 \leq k \leq \frac{n}{2} \\ 0 & \text{caso contrário} \end{cases} \quad \text{para } n \text{ par}$$

Se h é o filtro discreto cuja função de transferência é $H(\omega)$, então $(h_k)_{k \in \mathbb{Z}}$ é dada pela tabela 1 abaixo, para alguns valores de $n = \text{grau da B-spline}$.

k	n = 1	n = 2	n = 3
-2		0,125	0,0625
-1	0,5	0,375	0,25
0	0,	0,375	0,375
1	0,5	0,125	0,25
2			0,0625

Tabela 1: Coeficientes da função de transferência $H(\omega)$

As funções B-spline e Gaussiana possuem uma relação intrínseca, que é o fato que as $B^n(x)$ ou suas transformadas de Fourier convergem para Gaussianas [29] quando n tende a infinito. Desse modo, ao se tomar as wavelets ψ^n como a derivada primeira ou segunda de uma B-spline, o modelo aqui descrito, de extração de bordas, herdará características dos modelos já clássicos de Canny [7] ou de Marr-Hildreth [50]. Para que isto ocorra, toma-se

$$\hat{G}(\omega) = -ie^{-i\frac{\omega}{2}} \text{sen} \frac{\omega}{2}.$$

A transformada de Fourier da wavelet resultante será

$$\hat{\psi}^n(\omega) = \hat{G}\left(\frac{\omega}{2}\right) \hat{\phi}\left(\frac{\omega}{2}\right) = -i \frac{\omega}{4} e^{\frac{-i\omega(1+k)}{4}} \left(\text{sinc} \frac{\omega}{4}\right)^{n+2} \quad (4.7)$$

donde ψ^n é a derivada primeira de uma B-spline de grau $n+1$ centrada em

$$t = \frac{1+k}{4}$$

Se g é o filtro discreto cuja função de transferência é $G(\omega)$, então $(g_k)_{k \in \mathbb{Z}}$ é da forma, $g_0 = -g_1 = 0.5$ e $g_k = 0$ para os demais $k \in \mathbb{Z}$.

A Condição de Admissibilidade para esta família de wavelets, obtidas numericamente a partir das expressões (4.7) e (3.2), está apresentada na tabela 2 para alguns valores de n .

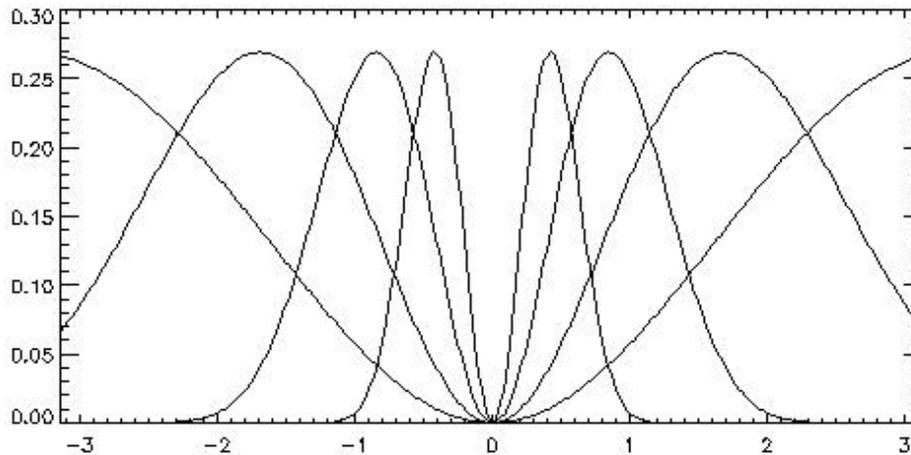


Figura 4.2: Gráfico do módulo da equação (4.7) para $0 \leq n \leq 3$

n	A	B
1	0,663	0,687
2	0,505	0,522
3	0,409	0,422
4	0,343	0,354

Tabela 2: Condição de Admissibilidade usando as fórmulas (3.2) e (4.7).

4.4 Implementação no Caso Discreto

Para o modelo discreto, assim como o modelo contínuo, a WT diádica de um sinal em $\ell^2(Z)$, isto é, sinais discretos de energia finita, é implementada como o operador linear limitado $\mathbf{W}: \ell^2(Z) \rightarrow \ell^2(I)$, onde $I = \{1, 2, \dots, J+1\} \times Z$, consistindo de $J+1$ operadores lineares $W_j: \ell^2(Z) \rightarrow \ell^2(Z)$, $j=1, 2, \dots, J+1$. Os operadores W_j são operadores de convolução com filtros cujas respostas a impulso são dadas por h_k e g_k .

Se $\hat{G}(\omega) = \sum g_n e^{-i\omega n}$, então $\hat{G}(2\omega) = \sum g_n e^{-i\omega 2n}$, o que corresponde a colocar um zero entre cada elemento do filtro discreto $g = \{g_n\}$. Mais geralmente, na para o nível p , coloca-se $2^p - 1$ zeros entre cada coeficiente do filtro g .

Sejam H_p e G_p os filtros discretos obtidos ao colocar-se $2^p - 1$ zeros entre cada coeficiente dos filtros H e G . Então, pelas relações de escala (4.5) e (4.6) obtém-se

$$\hat{W}_{2^j} f(\omega) = \hat{f}(\omega) \hat{\psi}(2^j \omega) = \hat{f}(\omega) \hat{G}(2^{j-1} \omega) \hat{\phi}(2^{j-1} \omega)$$

e, definindo-se $S_0 f = f$, a decomposição wavelet diádica de um sinal discreto f será implementada da seguinte forma:

$$W_{j+1} f = S_j f * G_j$$

$$S_{j+1} f = S_j f * H_j$$

onde, para efeito de simplificação, utiliza-se j ao invés de 2^j , isto é, $W_j f = W_{2^j} f$, $S_{2^j} f = S_j f$ $0 \leq j \leq J$ e $(*)$ denota convolução discreta. Uma representação esquemática é apresentada na figura 4.3.

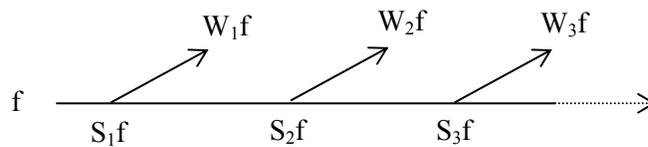


Figura 4.3: Esquema de decomposição 1D pela WT diádica

Se o sinal original $\{S_0 f(n); n \in \mathbb{Z}\}$ tem N amostras, é possível computar no máximo $\log_2(N)$ escalas da transformada Wavelet, visto que acima dessa escala, o filtro irá se tornar maior que o sinal.

4.5 Algoritmo Discreto Para Reconstrução

O procedimento descrito na figura 4.3, corresponde à parte da decomposição dada na figura 4.4. Em muitas situações, tais como em síntese de texturas, compressão de imagens, etc, é necessário que se tenha uma transformada inversa. A fim de tornar o procedimento de decomposição inversível, isto é, reconstruir $S_{j-1}f$ a partir de $S_j f$ e $W_j f$, deve-se construir filtros de síntese, \tilde{H} e \tilde{G} , conforme a parte de reconstrução da figura 4.4.

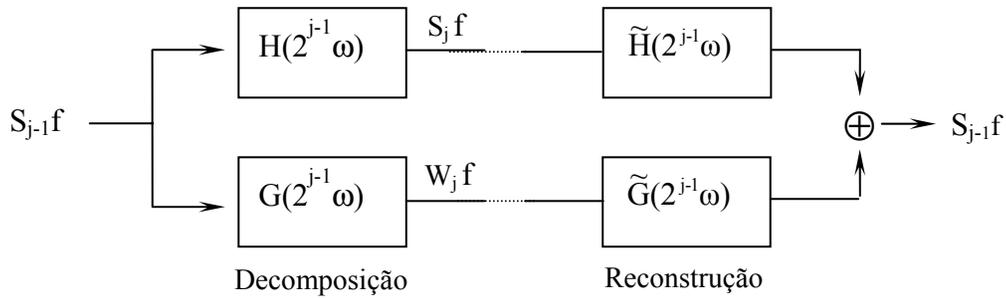


Fig. 4.4- Banco de filtros para a decomposição e reconstrução pela transformada wavelet

A saída, no espaço de freqüências desse sistema, será

$$\hat{S}_{2^{j-1}} f(\omega) \left[H(2^{j-1}\omega) \tilde{H}(2^{j-1}\omega) + G(2^{j-1}\omega) \tilde{G}(2^{j-1}\omega) \right].$$

Para assegurar que a saída do sistema seja idêntica a sua entrada, deve valer a seguinte identidade,

$$H(2^{j-1}\omega) \tilde{H}(2^{j-1}\omega) + G(2^{j-1}\omega) \tilde{G}(2^{j-1}\omega) = 1 \quad (4.8)$$

Tomando $\tilde{H} = \bar{H}$, isto é, o complexo conjugado de H , então \tilde{G} pode ser determinado unicamente por

$$\tilde{G}(2^{j-1}\omega) = \frac{1 - |H(2^{j-1}\omega)|^2}{G(2^{j-1}\omega)}$$

e a wavelet de reconstrução χ^n é definida como

$$\chi^n(2\omega) = \tilde{G}(\omega)\hat{\beta}(\omega)$$

donde f será reconstruída a partir de sua transformada wavelet diádica $W_j f$, por

$$f(x) = \sum_{j=-\infty}^{\infty} W_j f * \chi_j(x).$$

Pode-se obter \tilde{G} de modo explícito para as duas situações:

i- Para ψ definida conforme a equação (4.3),

$$\begin{aligned} \tilde{G}(\omega) &= \frac{1 - |H(\omega)|^2}{G(\omega)} = \frac{1 - \cos^{2n-2}\left(\frac{\omega}{2}\right)}{4i \operatorname{sen} \frac{\omega}{2} e^{i\frac{\omega}{2}}} = \frac{(1 - \cos^2 \frac{\omega}{2}) \sum_{j=0}^n (\cos^2 \frac{\omega}{2})^j}{4i \operatorname{sen} \frac{\omega}{2} e^{i\frac{\omega}{2}}} \\ &= \frac{1}{4i} \operatorname{sen} \frac{\omega}{2} e^{-i\frac{\omega}{2}} \sum_{j=0}^n (\cos^2 \frac{\omega}{2})^j \end{aligned}$$

A FIR correspondente é dada por:

$$\hat{g}_k = \begin{cases} \frac{1}{2^{2n+3}}, & k = -(n+1) \\ -\sum_{j=-k}^n \frac{1}{2^{2j+3}} \binom{2j}{j+k} + \sum_{j=-k}^{n+1} \frac{1}{2^{2j+1}} \binom{2(j-1)}{j+k}, & -n \leq k \leq -1 \\ -\sum_{j=-k}^n \frac{1}{2^{2j+3}} \binom{2j}{j+k} + \sum_{j=k}^{n+1} \frac{1}{2^{2j+5}} \binom{2(j+1)}{j+k+2}, & 0 \leq k \leq n-1 \\ -\frac{1}{2^{2n+3}}, & k=n \\ 0 & \text{caso contrário} \end{cases}$$

ii- Para ψ definida conforme a equação (4.4),

$$\tilde{G}(\omega) = \frac{1 - |H(\omega)|^2}{G(\omega)} = \frac{1 - \cos^{2n-2}\left(\frac{\omega}{2}\right)}{16 \operatorname{sen}^2 \frac{\omega}{2}} = -\frac{1}{16} \sum_{j=0}^n (\cos^2 \frac{\omega}{2})^j,$$

A FIR correspondente é dada por:

$$\tilde{g}_k = \begin{cases} -\frac{1}{16} \sum_{j=|k|}^n \frac{1}{2^{2j}} \binom{2j}{j+|k|}, & |k| \leq n, \\ 0 & \text{caso contrário.} \end{cases}$$

A partir da equação (4.8), obtém-se o algoritmo discreto para a reconstrução:

$$S_{2^{j-1}} f = S_{2^j} * \tilde{H}_{2^j} + W_{2^j} * \tilde{G}_{2^j}, \quad 0 \leq j \leq J$$

onde (*) denota convolução.

4.6 Transformada Wavelet Diádica de Imagens

A WT diádica pode ser estendida a $L^2(\mathbb{R}^2)$ introduzindo-se K wavelets, $\psi^k(x,y) \in L^2(\mathbb{R}^2)$, $1 \leq k \leq K$, onde cada wavelet $\psi^k(x,y) \in L^2(\mathbb{R}^2)$ poderá ter uma orientação específica. Na figura 4.5 pode ser vista uma representação esquemática da WT diádica com seis direções.

Seja

$$\psi_{2^j}^i(x,y) = \frac{1}{2^{2j}} \psi^i\left(\frac{1}{2^j}(x,y)\right),$$

a WT diádica de $f(x,y)$ na escala 2^j , em um ponto (x,y) e para k orientações é definida como:

$$\mathbf{W}_j f = (W_{2^j}^1 f(x,y), \dots, W_{2^j}^k f(x,y))$$

e, de modo geral, a transformada wavelet diádica com K direções é dada por

$$\mathbf{W}_k = [\mathbf{W}_j f], j \in \mathbb{Z}.$$

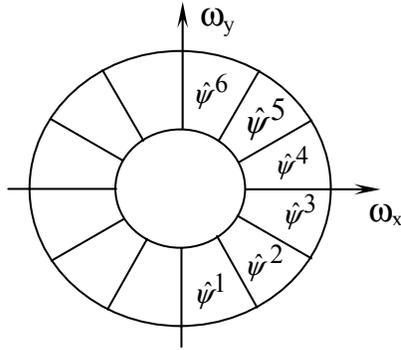


Figura 4.5: Wavelets com seis orientações

Para assegurar que exista uma representação completa e estável de $f(x, y)$, o plano de Fourier deve ser coberto pelas dilatações de $\hat{\psi}^k(\omega_x, \omega_y)$ [46]. Isto significa que deve existir $A_2, B_2 > 0$ tais que

$$\forall (\omega_x, \omega_y) \in \mathbb{R}^2 - \{(0,0)\}, \quad A_2 \leq \sum_{k=1}^K \sum_{j=-\infty}^{+\infty} \left| \hat{\psi}^k(2^j \omega_x, 2^j \omega_y) \right|^2 \leq B_2$$

Então existem wavelets de reconstrução $\{\chi^k(x, y)\}_{1 \leq k \leq K}$ tais que

$$\sum_{k=1}^K \sum_{j=-\infty}^{+\infty} \hat{\psi}^k(2^j \omega_x, 2^j \omega_y) \hat{\chi}^k(2^j \omega_x, 2^j \omega_y) = 1$$

onde

$$f(x, y) = \sum_{k=1}^K \sum_{j=-\infty}^{+\infty} (W_{2^j}^k f * \chi_{2^j}^k(x, y) + W_{2^j}^k f * \chi_{2^j}^k(x, y))$$

No modelo bidimensional, construído a seguir, serão utilizadas quatro orientações, $\theta = 0, 45, 90$ e 135 graus. Para isso, define-se as quatro wavelets ψ^1, ψ^2, ψ^3 e ψ^4 como

$$\psi^1(x, y) = \psi(x)2\phi(y),$$

$$\psi^2(x, y) = 2\phi(x)\psi(y),$$

$$\psi^3(x,y) = \psi^1\left(\frac{\sqrt{2}}{2}R_{45}(x,y)\right) = \psi^1\left(\frac{\sqrt{2}}{2}(x-y), \frac{\sqrt{2}}{2}(x+y)\right) \quad (4.9)$$

$$\psi^4(x,y) = \psi^2\left(\frac{\sqrt{2}}{2}R_{45}(x,y)\right) = \psi^2\left(\frac{\sqrt{2}}{2}(x-y), \frac{\sqrt{2}}{2}(x+y)\right)$$

onde $R_{45}(x,y) = (x-y, x+y)$ é a rotação de 45 graus no sentido anti-horário.

As quatro wavelets acima possuem as seguintes relações de escala

$$\begin{aligned} \hat{\psi}^1(\omega_x, \omega_y) &= \hat{\psi}(\omega_x) \hat{\phi}\left(\frac{\omega_y}{2}\right) = G\left(\frac{\omega_x}{2}\right) \hat{\phi}\left(\frac{\omega_x}{2}\right) \hat{\phi}\left(\frac{\omega_y}{2}\right) \\ \hat{\psi}^2(\omega_x, \omega_y) &= \hat{\phi}\left(\frac{\omega_x}{2}\right) \hat{\psi}(\omega_y) = \hat{\phi}\left(\frac{\omega_x}{2}\right) G\left(\frac{\omega_y}{2}\right) \hat{\phi}\left(\frac{\omega_y}{2}\right) \\ \hat{\psi}^3(\omega_x, \omega_y) &= \hat{\psi}^1(\omega_x - \omega_y, \omega_x + \omega_y) = G\left(\frac{\omega_x - \omega_y}{2}\right) \hat{\phi}\left(\frac{\omega_x - \omega_y}{2}\right) \hat{\phi}\left(\frac{\omega_x + \omega_y}{2}\right) \\ \hat{\psi}^4(\omega_x, \omega_y) &= \hat{\psi}^2(\omega_x - \omega_y, \omega_x + \omega_y) = \hat{\phi}\left(\frac{\omega_x - \omega_y}{2}\right) G\left(\frac{\omega_y + \omega_x}{2}\right) \hat{\phi}\left(\frac{\omega_y + \omega_x}{2}\right) \end{aligned} \quad (4.10)$$

Se $G(\omega) = \sum g_k e^{-ik\omega}$, então

$$G(\omega_x - \omega_y) = \sum g_k e^{-ik(\omega_x - \omega_y)} = \sum g_k e^{-ik\omega_x} e^{ik\omega_y} \quad (4.11)$$

e os coeficientes de $G(\omega_x \pm \omega_y)$ são os mesmos de $G(\omega)$, mas atuam na diagonal da imagem, ou seja, $G(\omega_x \pm \omega_y)$ também é um filtro FIR.

4.7 Condição de Admissibilidade para ψ^1 , ψ^2 , ψ^3 , ψ^4

Já foi visto que, para a família de wavelets ψ^n , unidimensionais, construídas na seção 4.3, existem constantes A e B, com $0 \leq A, B$, tais que

$$A \leq \sum_{j=-\infty}^{+\infty} \left| \hat{\psi}(2^j \omega) \right|^2 \leq B$$

onde A e B são dados pela tabela 2, conforme o grau n da B-spline utilizada. Assim, conforme (3.2), deve-se mostrar que existem constantes A_k, B_k tais que

$$A_k \leq \sum_{k=1}^4 \sum_{j=-\infty}^{+\infty} \left| \hat{\psi}^k(2^j \omega_x, 2^j \omega_y) \right|^2 \leq B_k$$

onde ψ^k , $1 \leq k \leq 4$, são dadas por (4.9) e correspondem às direções especificadas. Para isso, considere inicialmente os cálculos realizados abaixo para a wavelet ψ^1 .

$$\sum_{j=-\infty}^{+\infty} \left| \hat{\psi}^1(2^j \omega_x, 2^j \omega_y) \right|^2 = \sum_{j=-\infty}^0 \left| \hat{\psi}^1(2^j \omega_x, 2^j \omega_y) \right|^2 + \sum_{j=1}^{+\infty} \left| \hat{\psi}^1(2^j \omega_x, 2^j \omega_y) \right|^2$$

mas,

$$\begin{aligned} \sum_{j=-\infty}^0 \left| \hat{\psi}^1(2^j \omega_x, 2^j \omega_y) \right|^2 &= \sum_{j=-\infty}^0 \left| \hat{\psi}(2^j \omega_x) \hat{\phi}(2^{j-1} \omega_y) \right|^2 = \sum_{j=-\infty}^0 \left| \hat{\psi}(2^j \omega_x) \right|^2 \left| \hat{\phi}(2^{j-1} \omega_y) \right|^2 \leq \\ &\leq \sum_{j=-\infty}^0 \left| \hat{\psi}(2^j \omega_x) \right|^2 \end{aligned}$$

$$\text{pois } \left| \hat{\phi}(2^{j-1} \omega_y) \right|^2 = \left| e^{-ik\frac{\omega}{2}} \left(\frac{\text{sen} \frac{\omega}{2}}{\frac{\omega}{2}} \right)^{n+1} \right|^2 \leq 1. \quad (4.12)$$

Para os valores positivos de j na série, segue que

$$\lim_{j \rightarrow \infty} \left| \hat{\phi}(2^{j-1} \omega_y) \right|^2 = \lim_{j \rightarrow \infty} \left| \frac{\text{sen}(2^{j-1} \omega_y)}{2^{j-1} \omega_y} \right|^2 = 0 \quad (4.13)$$

e como $\sum_{j=0}^{\infty} \left| \hat{\psi}(2^j \omega_x) \right|^2$ converge (caso unidimensional), então, pelo teste de Dirichlet para séries, a série do produto dos módulos converge. Assim, (4.12) e (4.13) são suficientes para garantir que existem constantes A_k e B_k , tais que

$$A_k \leq \sum_{j=-\infty}^{+\infty} \left| \hat{\psi}^1(2^j \omega_x, 2^j \omega_y) \right|^2 \leq B_k$$

O caso ψ^2 é análogo ao caso ψ^1 . Quanto a ψ^3 e ψ^4 , basta notar que $\hat{\psi}(\omega_x \pm \omega_y) = \omega_z$ para algum z , e, portanto, fica garantida a existência das constantes A e B .

4.8 Algoritmo Discreto

As relações de escala em (4.10), juntamente com (4.11), permitem uma implementação rápida da transformada wavelet diádica. Para isto, seja $D_1 F$ a matriz cuja diagonal principal é formada pelos coeficientes do filtro F (aqui F é tomado como um filtro qualquer); $D_2 F$ a matriz cuja diagonal secundária é formada pelos coeficientes do filtro F ; e F_p o filtro discreto obtido ao colocar-se $2^p - 1$ zeros entre cada coeficiente do filtro F . Desse modo, a decomposição diádica discreta de um sinal bidimensional f é feita da seguinte forma:

$$W_{2^{j+1}}^1 f = S_{2^j}^1 f * (G_j \cdot I)$$

$$W_{2^{j+1}}^2 f = S_{2^j}^1 f * (I \cdot G_j)$$

$$W_{2^{j+1}}^3 f = S_{2^j}^1 f * D_1 G_j$$

$$W_{2^{j+1}}^4 f = S_{2^j}^1 f * D_2 G_j$$

$$S_{2^{j+1}}^1 f = S_{2^j}^1 f * (H_j \cdot I) * (I \cdot H_j)$$

$$S_{2^{j+1}}^2 f = S_{2^j}^2 f * (D_1 H_j) * (D_2 H_j)$$

onde (*) significa convolução e I é o filtro de Dirac, cuja resposta a impulso é igual a 1 na posição zero, e zero nas demais posições, cujo uso é apenas notacional, com o objetivo de tornar as convoluções bidimensionais; além disso, o uso dos índices 2^j expressam as escalas diádicas. Para exemplificar, se

$$G_0 = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ -\frac{1}{2} \\ -\frac{1}{2} \end{bmatrix}, \text{ então } G_1 = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ 0 \\ -\frac{1}{2} \\ -\frac{1}{2} \end{bmatrix}, \quad I = [1, 0, 0, \dots], \quad G_0 \cdot I = \begin{bmatrix} \frac{1}{2} & 0 \\ -\frac{1}{2} & 0 \end{bmatrix}, \quad I \cdot G_0 = \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ 0 & 0 \end{bmatrix}$$

$$D_1 G_0 = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & -\frac{1}{2} \end{bmatrix}, \quad D_2 G_0 = \begin{bmatrix} 0 & -\frac{1}{2} \\ \frac{1}{2} & 0 \end{bmatrix}, \quad D_1 G_1 = \begin{bmatrix} \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -\frac{1}{2} \end{bmatrix}.$$

Uma versão esquemática da implementação discreta do modelo de filtragem bidimensional é apresentada na figura 4.7.

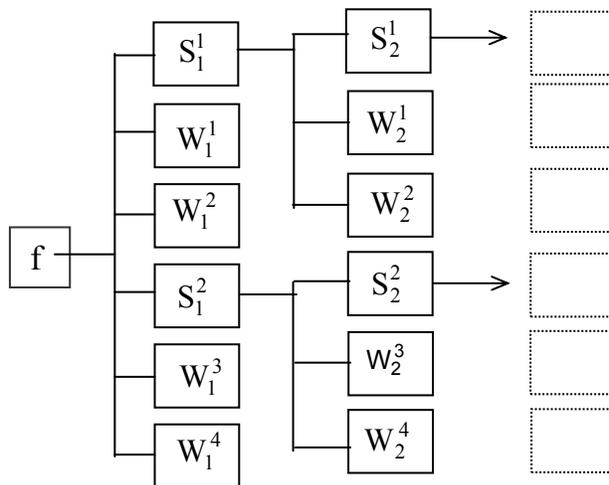


Figura 4.7: Esquema da decomposição de um sinal bidimensional pela transformada wavelet diádica com quatro direções.

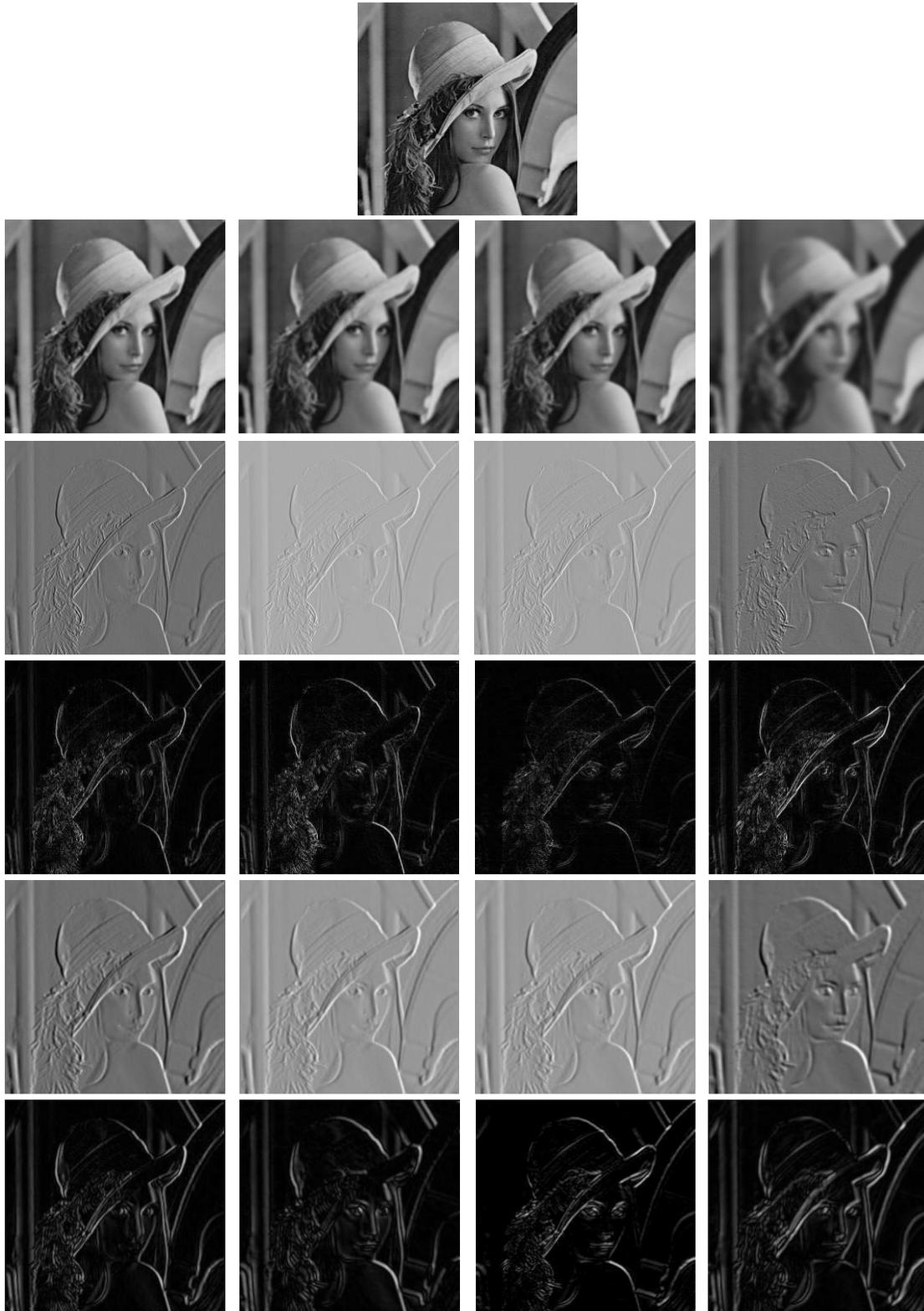


Figura x.y: Linha1: Imagem original; linha 2: $S_2^1, S_2^2, S_4^1, S_4^2$, linha 3: $W_2^1, W_2^2, W_2^3, W_2^4$, linha 4: $|W_2^1|, |W_2^2|, |W_2^3|, |W_2^4|$, linha 5: $W_4^1, W_4^2, W_4^3, W_4^4$, linha 6: $|W_4^1|, |W_4^2|, |W_4^3|, |W_4^4|$. Para fins de visualização, as imagens de média zero foram colocadas no range [0 255].

4.9 Sumário

Nesse capítulo, fez-se o desenvolvimento da fundamentação matemática teórica utilizada nesse trabalho. Construiu-se uma família de wavelets diádicas para a decomposição unidimensional à partir das B-spline. Para o caso de wavelets bidimensionais, construiu-se um modelo com quatro orientações e mostrou-se como se faz a implementação no caso discreto, além disso, apresentou-se um exemplo prático a fim de ilustrar a decomposição wavelet diádica com quatro direções em uma imagem.

Capítulo 5

Fundamentação e Modelo Experimental

5.1 Bordas em Multiescala

Conforme exposto no capítulo 1, a discriminação de texturas envolve identificar regiões com texturas uniformes em uma dada imagem, tendo em conta que textura é propriedade de uma região e não de um ponto.

Uma textura mais lisa ou mais rugosa está relacionada a variações tonais locais menos ou mais bruscas, ou seja, descontinuidades locais. Os pontos onde há intensas variações tonais estão entre as mais importantes características para analisar as propriedades de sinais transientes ou imagens. Em imagens, eles fornecem a localização das bordas, que geralmente são usadas para descrever a informação estrutural dos contornos de objetos. As bordas de estruturas em imagens são freqüentemente os mais importantes atributos para o reconhecimento de padrões. Isto é bem ilustrado pela capacidade visual dos seres humanos de reconhecer um objeto a partir de um esboço, grosseiro até, de seus contornos. Mas, o que é uma borda? Pode-se definir borda como pontos onde a intensidade numa imagem sofre uma variação brusca; contudo, quando se tratando de texturas em imagens, estas contêm variações bruscas de intensidade que freqüentemente não são consideradas como bordas. Por exemplo, olhando uma parede de tijolos, pode-se dizer que as bordas são os contornos da parede, enquanto os tijolos definem a textura, figura 5.1(a). Da mesma forma, pode-se

dizer que o contorno de um tijolo é uma borda, enquanto que a rugosidade de sua superfície determina a textura, figura 5.1(b). Assim, a discriminação de bordas ou texturas depende da escala de análise e estão intrinsecamente relacionadas.

Este fato motivou pesquisadores da área de visão computacional a buscarem técnicas de detectar variações abruptas em imagens em diferentes escalas, [50], [44].

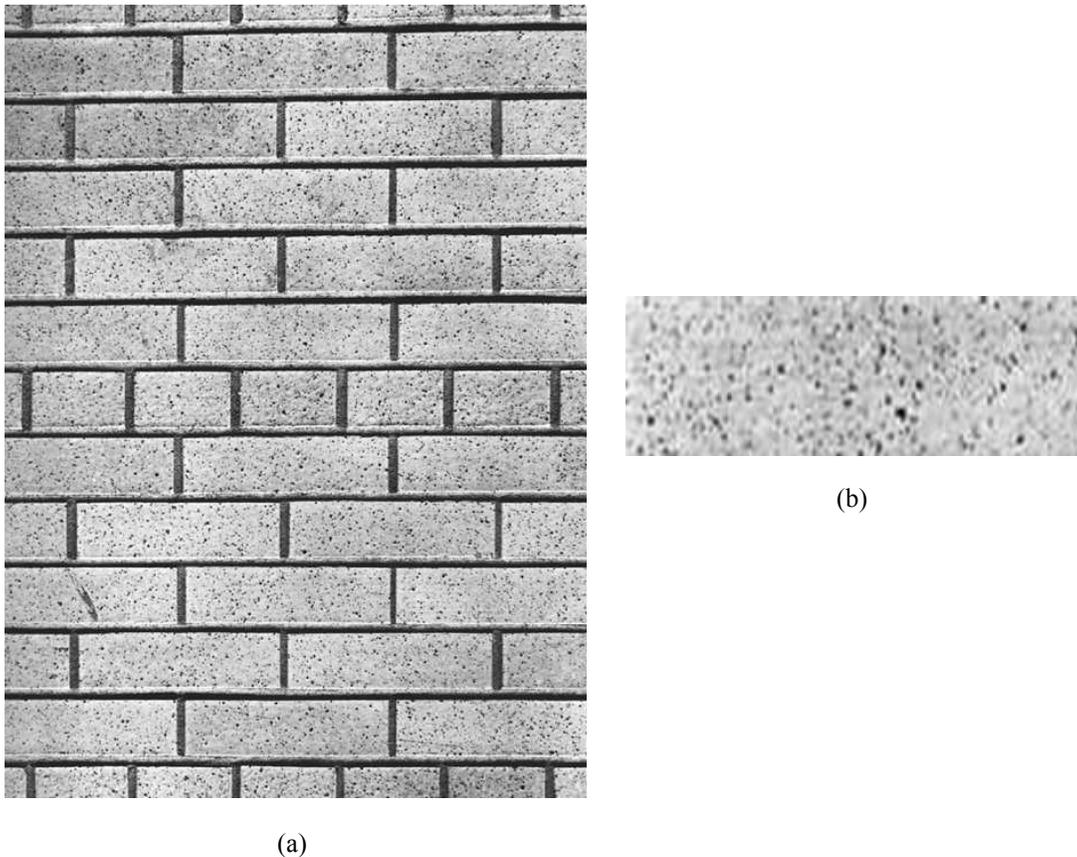


Figura 5.1: (a) textura formada por tijolos; (b) textura da superfície de um único tijolo.

Na maioria das vezes, detecção de bordas é visto como um processo que reduz a quantidade de dados representando o sinal, mas não preservando toda a informação; contudo, Mallat et al [44], [45] mostraram que bordas em multiescala da transformada wavelet fornecem uma representação completa e estável do sinal. Nos trabalhos de Mallat, são consideradas bordas, o módulo das saídas da transformada wavelet diádica.

Nesse trabalho, a discriminação de texturas é obtida a partir da quantificação do conteúdo de bordas (energia), detectado pela transformada wavelet diádica de imagens em suas várias escalas e orientações, através da análise da energia local nessas escalas. Tal análise é precedida de algumas transformações, lineares e não lineares nas saídas da transformada wavelet.

Essa abordagem se insere no contexto conhecido como filtragem por multi-canais, inspirada na teoria da filtragem por multi-canais, usada para processar informação visual nos estágios primários (*early stages*) do sistema visual humano.

A partir de estudos da fisiologia da visão humana, foram encontradas evidências mostrando que a imagem na retina é decomposta em vários canais de frequência espacialmente orientados. Esta teoria foi proposta por Campbell e Robson [6]. Experimentos psicofisiológicos posteriores forneceram evidências adicionais como suporte a essa teoria [20], sugerindo que existem mecanismos no córtex visual dos mamíferos que são sintonizados (*tuned*) a combinações de frequência e orientação em um range limitado (*narrow range*). Esses mecanismos são freqüentemente referidos como canais, e interpretados como filtros passa-banda.

A fim de imitar esse procedimento visual, desenvolveu-se um modelo de decomposição bidimensional com quatro orientações por nível da transformada wavelet, onde os níveis e as orientações são os equivalentes aos canais de frequência; além disso, em termos matemáticos, essa é uma representação completa e estável de um sinal discreto bidimensional, isto é, a energia do sinal é preservada e é possível reconstruí-lo.

5.2 Multirresolução e Wavelets

Técnicas utilizando multirresolução para processamento de imagens e sinais têm sido utilizadas há décadas [1]. A necessidade por transformações que levam a uma representação em que as informações espaciais e de frequências estão presentes, tem sua importância reconhecida. A fim de obter tal

representação, uma grande variedade de técnicas foram desenvolvidas, tais como, Gabor [28], Haar [31], Walsh-Hadamard [84] e outras expansões, além de filtragem por sub-bandas, decomposição espaço-escala, etc.

Por volta de 1990, surge a primeira ferramenta matemática que fornece uma abordagem mais formal, sólida e unificada, para uma representação em multirresolução [47], consistindo de uma decomposição multirresolução de sinais em bases ortogonais de funções wavelets de suporte compacto (zero fora de um intervalo fechado). O desenvolvimento de métodos para construir tais funções mostrou-se um aspecto importante da teoria. De fato, o advento de wavelets suaves e compactas deu início à inovação desse ferramental, pois tornou possível computar decomposições precisas com limitado esforço computacional. Isso propiciou o surgimento, em poucos anos, de muitas aplicações em processamento de sinais e imagens. Além disso, muitas teorias precursoras, como por exemplo o trabalho de Burt e Adelson [4] em visão computacional e o trabalho de Croisier, Esteban e Galand [16] utilizando banco de filtros para compressão de sons, puderam ser reformuladas em terminologia wavelet, tornando-a uma técnica padrão para processamento de sinais em multirresolução.

5.3 Modelo Experimental

A asserção básica feita nesse trabalho é que medidas da distribuição da energia local, resultante da decomposição do sinal dada pela transformada wavelet diádica discreta nas suas várias escalas e orientações, são suficientes para discriminar inúmeras texturas. Isto decorre do fato que o módulo da transformada wavelet, $|W^k f(u, v, 2^j)|^2$, é máximo quando o ângulo α_k de W^k coincide com a orientação dos componentes da textura e quando 2^j corresponde à escala dessas estruturas. Veja figura 5.2.

O uso da energia local tem sido bastante utilizado não só no contexto da transformada wavelet, mas também no da transformada de Gabor ou mesmo nos de matriz de co-ocorrência, [35], [58], [81].

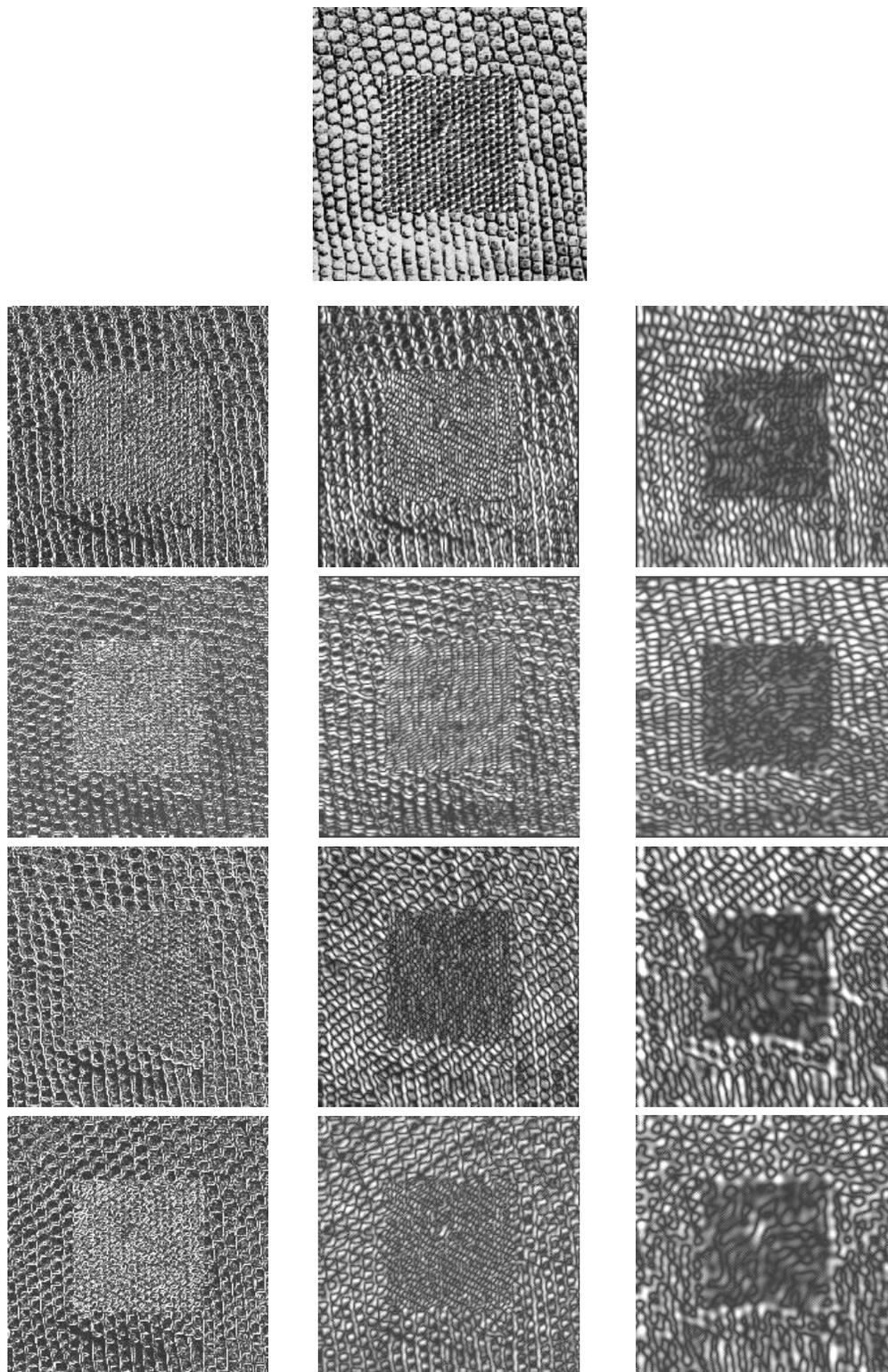


Figura 5.2: Módulo da transformada wavelet, $\left| W^k f(u, v, 2^j) \right|^2$, nos níveis 1 (coluna esquerda), 2 (coluna central) e 3 (coluna direita) e nas direções, de cima para baixo, 0° , 45° , 90° e 135° graus.

Para fins de segmentação, a principal dificuldade é encontrar um algoritmo que agregue as respostas da transformada wavelet em todas as escalas e orientações de modo a encontrar as bordas de regiões de texturas homogêneas. Dentro de uma única região, para cada escala e orientação, a energia da wavelet, $|W^k f(u, v, 2^j)|^2$, pode ter um alto grau de variabilidade, como mostrado pela figura 5.2. Nesse caso, pode-se notar visualmente que no nível três, sem qualquer processamento posterior, a transformada wavelet do par de texturas apresenta um bom contraste visual, e, no nível dois da orientação de 90° , essa boa separabilidade também já foi obtida. Entretanto, para outros níveis e orientações, o resultado não é tão bom, podendo induzir erro no resultado final. Assim, como o objetivo é obter a separabilidade automática, deve-se utilizar algum procedimento que selecione as melhores bandas.

O modelo esquemático experimental apresentado na figura 5.3 descreve o procedimento utilizado, no intuito de realizar e otimizar esta tarefa. Esse esquema ilustra as “operações” realizadas no nível J da transformada wavelet, que são: a decomposição wavelet diádica, a tomada do valor absoluto das saídas da transformada wavelet, o cálculo das energias locais após uma homogeneização, a combinação de atributos e, por fim, a seleção de uma banda ótima de melhor separabilidade. Todo esse procedimento, descrito de modo simplificado nesse parágrafo, encontra-se explicitado na próxima seção.

5.4 Etapas do Modelo Experimental

A fim de melhor compreender as etapas do modelo experimental, é interessante que o leitor se reporte às figuras 5.3 e 5.4 quando da descrição das etapas. Todo esse procedimento é realizado em um par de amostras tomadas em duas texturas distintas. Os parâmetros obtidos nesse treinamento serão utilizados na imagem original a fim de segmentar os pares de texturas. Para o caso de mosaicos com mais de duas texturas, aplica-se o método combinatorial descrito na seção 5.6. Esse método funciona bem se a discriminação de pares de texturas

estiver funcionando adequadamente. O usuário, poderia também utilizar um classificador de sua preferência. A figura 5.3 descreve o procedimento num nível J da transformada e obviamente representa o que ocorre nos outros níveis.

Na primeira etapa do modelo, a amostra de textura é convolucionada com um banco de quatro filtros, isto é, faz-se a decomposição wavelet diádica nas quatro orientações. Neste estágio, as informações correspondentes a diferenças entre regiões de texturas estão concentradas em diferenças nas variâncias dos atributos iniciais, visto que todas as bandas possuem inicialmente médias zero. A segunda etapa consiste em retificar e homogeneizar as respostas dos filtros. A retificação é feita tomando-se o valor absoluto das saídas da transformada e corresponde a uma transformação não linear. Esta operação tem o efeito de converter diferenças de variância em diferenças de médias a fim de trabalhar com a energia resultante nas bandas obtidas. Com relação à homogeneização subsequente nas bandas, ela se insere no contexto de que textura é propriedade local e não pontual, ou seja, não é definida para um único pixel, mas está associada a um conjunto de pixels. Ela possibilita regiões de texturas mais homogêneas, devido à redução da variância das energias, resultando em classes no espaço atributo que são mais compactas e conseqüentemente mais fáceis de distinguir.

A homogeneização pode ser feita com vários tipos de filtros. Alguns autores, [71] [85], sugerem o filtro Gaussiano passa baixa,

$$h_{\sigma}(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{x^2}{2\sigma^2}\right).$$

Uma versão discreta, baseada no Teorema Central do Limite [29], que diz que a gaussiana é o limite de convoluções sucessivas do filtro caixa⁴ (*box filter*) é a que se utilizou nesse trabalho. Uma discretização unidimensional do filtro gaussiano por aproximações sucessivas, pode ser feita tomando-se o filtro box e considerando-se suas convoluções:

$$\text{box} = \frac{1}{2} [1 \quad 1]; \quad \text{box} * \text{box} = \frac{1}{4} [1 \quad 2 \quad 1]; \quad \text{box} * \text{box} * \text{box} = \frac{1}{8} [1 \quad 3 \quad 3 \quad 1].$$

⁴ Box: filtro que realiza uma média simples em região quadrangular.

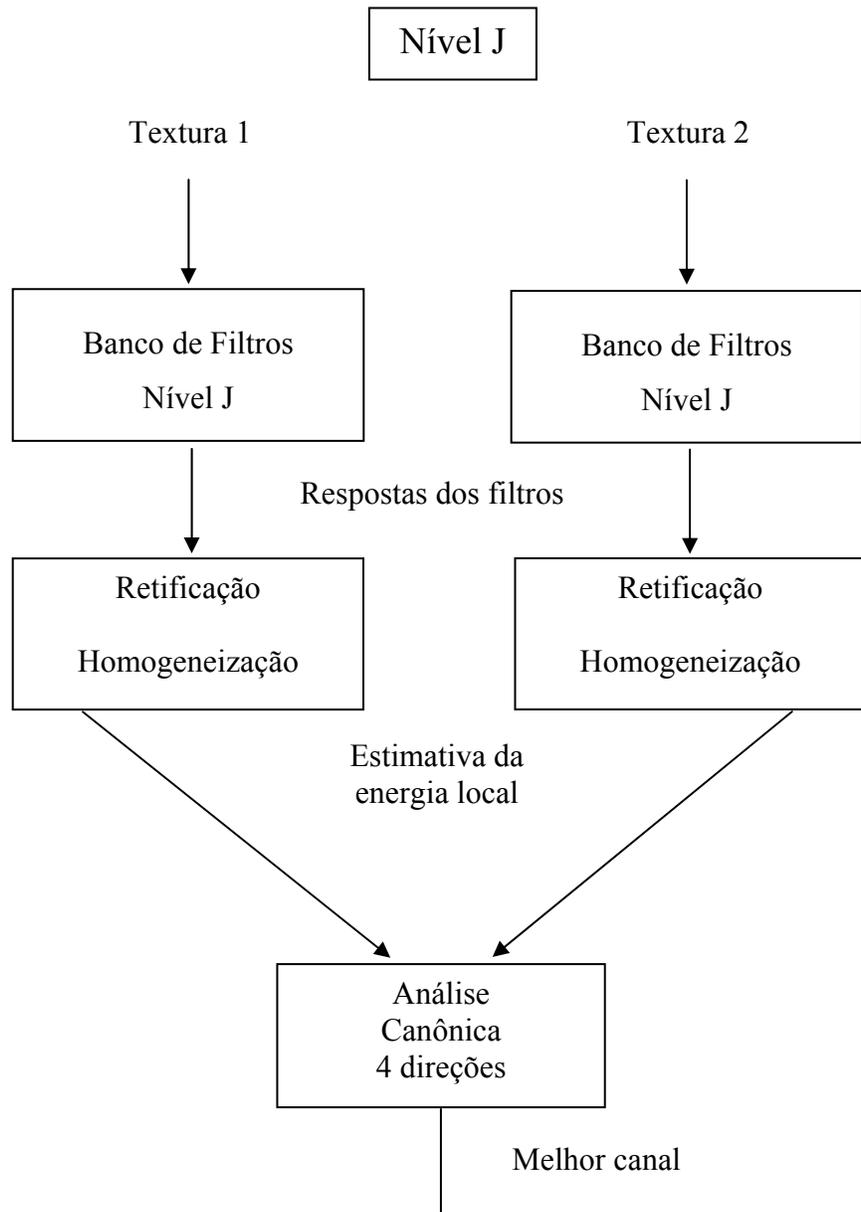


Figura 5.3: Esquema representativo do modelo experimental, etapas 1, 2, e 3.

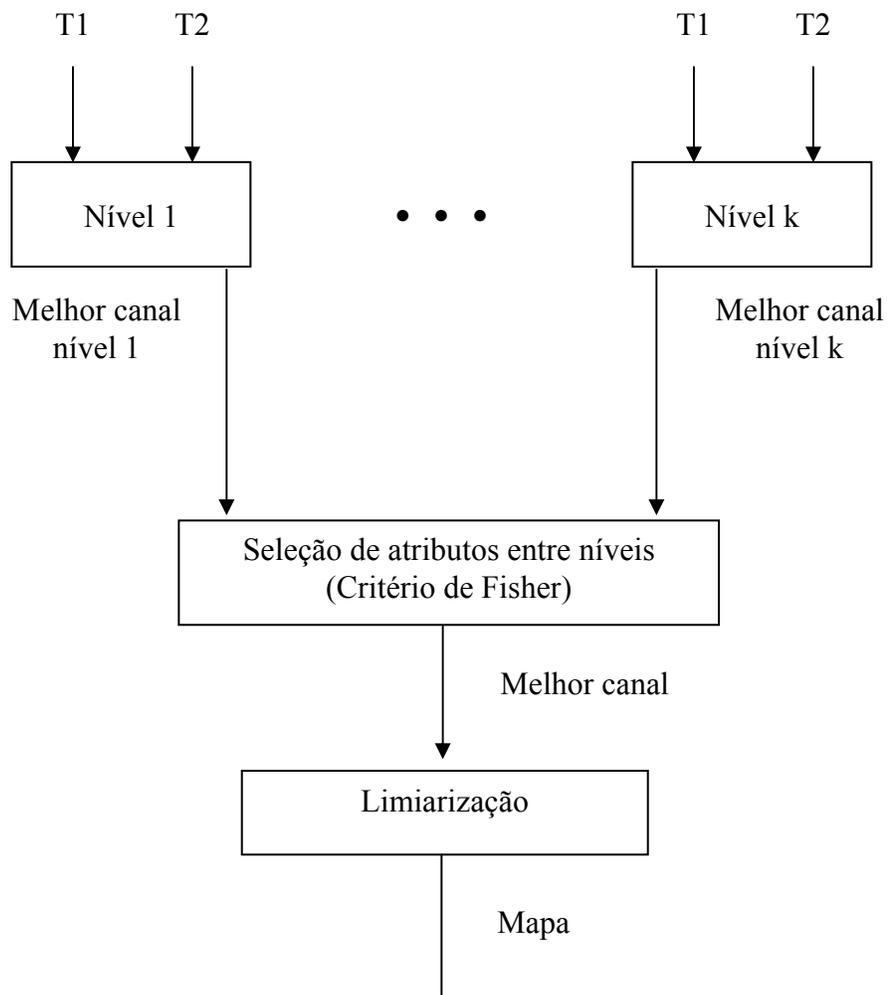


Figura 5.4: Esquema representativo do modelo experimental, etapas 4 e 5.

De modo geral, a máscara unidimensional de ordem $k+1$ é dada por:

$$\frac{1}{2^k} [1 \ k \ \dots \ \binom{k}{i-1} \binom{k}{i} \binom{k}{i+1} \ \dots \ 1]$$

No caso bidimensional, usa-se a Gaussiana separável, isto é,

$$h_\sigma(x,y) = \frac{1}{\sigma^2 2\pi} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) = h_\sigma(x)h_\sigma(y).$$

Essa etapa pode ser considerada como uma estimativa da energia local. O efeito prático dessa etapa pode ser visto na figura 5.5.

Após a homogeneização, obtêm-se quatro bandas por nível da transformada Wavelet. Cada banda traz consigo alguma informação sobre a imagem. Apesar da redundância inerente a essa decomposição, cada banda possui alguma informação diferenciada em relação à outra, visto que são oriundas de orientações específicas.

Na terceira etapa, a fim de se obter a maior separabilidade possível entre pares de texturas, aplica-se a Análise Canônica [72], que opera da seguinte forma: para M classes e N bandas de dados multiespectrais, existem $M-1$ eixos canônicos. Em nosso caso, $N=4$ e $M=2$, logo, existe uma projeção num único eixo que fornece a máxima separação entre as classes. Mais detalhadamente, para as bandas B_1, B_2, B_3 e B_4 geradas na etapa 2, existe um vetor $d = [\alpha_1, \alpha_2, \alpha_3, \alpha_4]$, tal que a banda

$$B_j = \alpha_1 B_1 + \alpha_2 B_2 + \alpha_3 B_3 + \alpha_4 B_4$$

apresenta a melhor separabilidade no nível J . No apêndice desse trabalho, encontra-se um programa em IDL, chamado "Canonic4.pro" que implementa a análise canônica para quatro bandas e duas classes.

Na quarta etapa, tem-se em mãos uma banda de máxima separabilidade para cada nível da transformada. A fim de escolher qual o melhor deles, usa-se o critério de separabilidade de Fisher em cada banda de melhor separabilidade,

$$C_F = \frac{(\mu_1 - \mu_2)^2}{\sigma_1^2 + \sigma_2^2}$$

onde μ_1 e σ_1 e são a média e a variância da amostra da textura 1 e μ_2 e σ_2 são a média e a variância da amostra da textura 2, tomadas agora nas bandas de máxima separabilidade. Note que esse critério leva em conta não apenas a separação das médias das energias, mas também o quanto essas energias se afastam das médias, isto é, as variâncias. Daquela que obtém o maior índice no Critério de Fisher, toma-se seus parâmetros, que são o nível e os coeficientes do vetor d , e aplica-se na imagem original para gerar a banda de máxima separabilidade entre as texturas 1 e 2. A escolha do nível depende da experiência do usuário e principalmente das escala dos objetos de sua aplicação.

Na etapa final, faz-se uma limiarização na banda de máxima separabilidade para se obter a segmentação da imagem. Como textura é característica de uma região e não de pixels isolados, uma forma automática de obter essa limiarização 'pode ser feita tomando-se as energias nas mesmas regiões que se colheu as amostras e usá-las como centro de classe. Para cada pixel, tomar a energia local e verificar a qual centro de classe ele está mais próximo, classificando-o nessa classe.

Resumidamente, para cada par de texturas T_1 e T_2 a serem segmentadas, realiza-se o seguinte procedimento:

- 1- Tomam-se amostras representativas de cada textura e, para cada nível da transformada wavelet diádica, obtêm-se quatro bandas, b_1 , b_2 , b_3 e b_4 , utilizando os filtros gerados pela implementação discreta da transformada wavelet.
- 2- Calcula-se o valor absoluto.
- 3- Aplica-se o filtro gaussiano.
- 4- A partir dessas quatro bandas, realiza-se a Análise Canônica, obtendo-se uma banda para cada nível.
- 5- Compara-se a discriminabilidade entre os níveis, utilizando o critério de Fisher, a fim de verificar qual a que possui maior discriminabilidade, obtendo-

se os parâmetros finais que serão utilizados na imagem original. Os parâmetros finais consistem do nível da transformada e dos coeficientes da rotação canônica.

- 6- Faz-se a segmentação das duas texturas. Para isso, utilizam-se como centro de classes os valores médios das energias obtidas na posição das amostras da banda que mostrou máxima separabilidade.

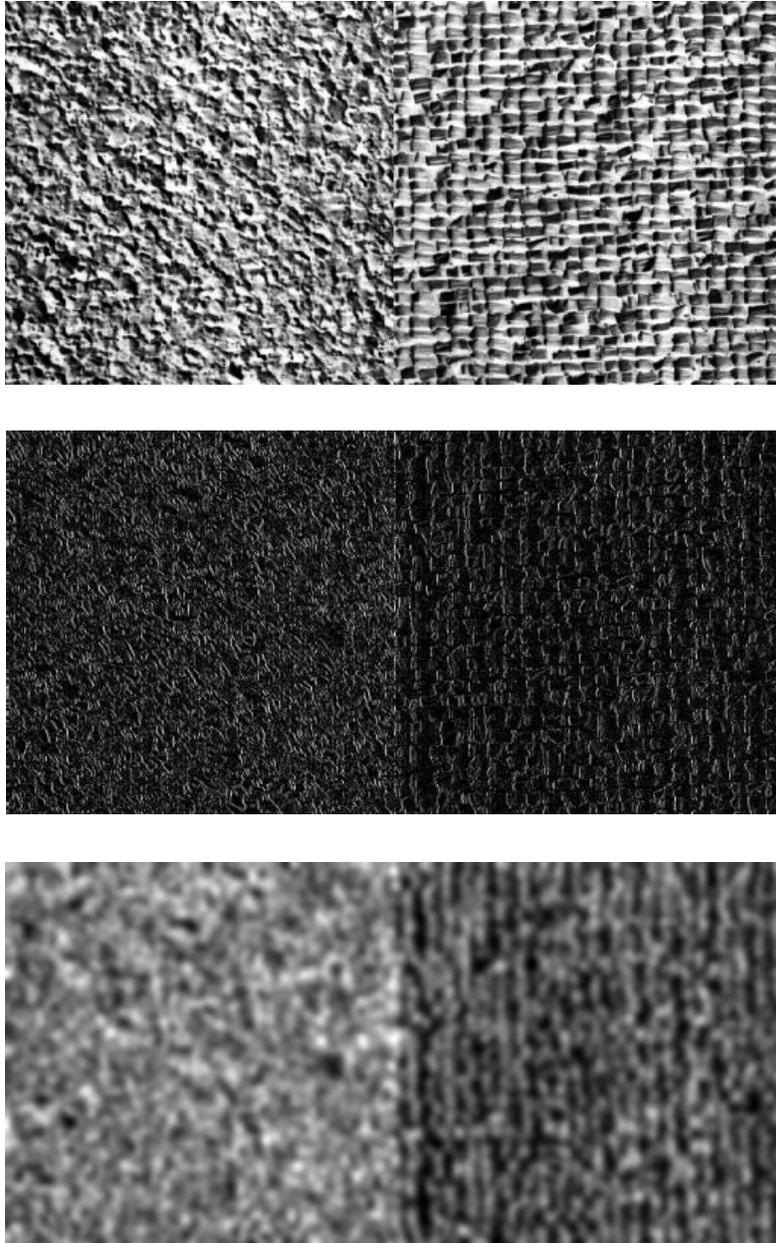


Figura 5.5: Superior: Texturas D04D084 de Brodatz [3], Meio: Não linearidade, Inferior: Imagem homogeneizada.

A significância de estudar pares de texturas vem do fato que todo problema de análise de texturas pode ser dividido em um problema equivalente em que se considera a discriminação de pares de texturas, figura 5.6.

No caso de mais que duas texturas, pode-se aplicar o mesmo procedimento que se usa para pares de texturas e em seguida usar alguma técnica de segmentação. Nesse trabalho, utiliza-se um algoritmo combinatorial, que, além de ser facilmente implementado, apresenta bons resultados.

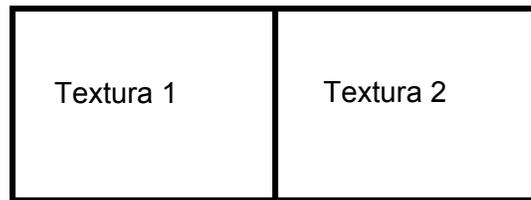


Figura 5.6: Modelo de par de texturas.

5.5 Algoritmo Combinatorial

O problema de segmentar mais que duas texturas pode ser resolvido usando-se o método para a segmentação de pares de texturas, da seguinte forma: para cada par de texturas, é gerada uma banda que discrimina maximamente este par, num treinamento com amostras dessas duas texturas. Aplicando-se o resultado desse treinamento numa imagem composta por n texturas, obtém-se

$\binom{n}{2} = \frac{n(n-1)}{2}$ bandas. Assim, para cada pixel da imagem, obtém-se um vetor

atributo de $\frac{n(n-1)}{2}$ posições, onde, na posição 1 terá o valor 1 ou 2, caso o pixel

tenha sido classificado como pertencendo à textura 1 ou à textura 2, na banda 1.

Na posição 2, o vetor terá o valor 1 ou 3, caso o pixel tenha sido classificado como pertencendo à textura 1 ou 2, na banda 2. Assim, sucessivamente, na posição

$\frac{n(n-1)}{2}$ o vetor atributo terá o valor $n-1$ ou n , conforme o pixel tenha sido

classificado como pertencendo à textura $n-1$ ou à textura n , na banda $n(n-1)/2$. Por exemplo, considere um pixel pertencendo à textura 1, num mosaico de 4 texturas. Tipicamente, seu vetor atributo será da forma $(1,1,1,x,y,z)$. Note que os valores x , y e z , não podem ser iguais, pois, $x \in \{2, 3\}$, $y \in \{2, 4\}$ e $z \in \{3, 4\}$. Assim, basta verificar em qual classe o pixel é classificado mais vezes e associá-lo a essa classe. Esse algoritmo não é custoso computacionalmente, pois o treinamento a fim de se obter os parâmetros é realizado em amostras das texturas. Com essas amostras, gera-se a banda ótima, faz-se a limiarização e obtém-se o valor de entrada para o vetor atributo, podendo-se descartar a banda ótima a fim de refazer o procedimento para um novo par de texturas, não havendo a necessidade de manter na memória da máquina as bandas ótimas geradas.

Para exemplificar, considere o mosaico de três texturas, T1, T2, e T3 da figura 5.6. A textura T1 forma o centro do mosaico, a textura T2 está na parte inferior do mosaico e a textura T3 na parte superior do mosaico. Pelo algoritmo combinatorial, a cada pixel da imagem original estará associado um vetor atributo de três posições cujos valores variam de 1 a 3. Os pixels serão classificados como pertencendo às texturas que tiverem mais entradas nos vetores atributo associados a eles. A figura 5.7 tem dimensões 256x256 pixels. Vejamos nessa figura, duas situações distintas, uma em que a segmentação foi executada corretamente e a outra em que isso não ocorreu. Considere o pixel P_1 na posição (128,128). Na figura 5.7(d), esse pixel foi, pelo limiar, classificado como pertencendo à textura 1, em 5.7(e) como pertencente à textura 1 e em 5.7(f) como pertencendo à textura 2. Assim, seu vetor atributo é da forma (1,1,2) e portanto foi classificado como pertencendo à textura 1 na segmentação final. Esse é um exemplo que deu certo. Consideremos agora um pixel na mancha preta no canto inferior direito da figura 5.7(c). Aquela região foi classificada erroneamente. Um pixel nessa região possui vetor atributo da forma (1,1,2) e portanto foi classificado como pertencendo à textura 1.

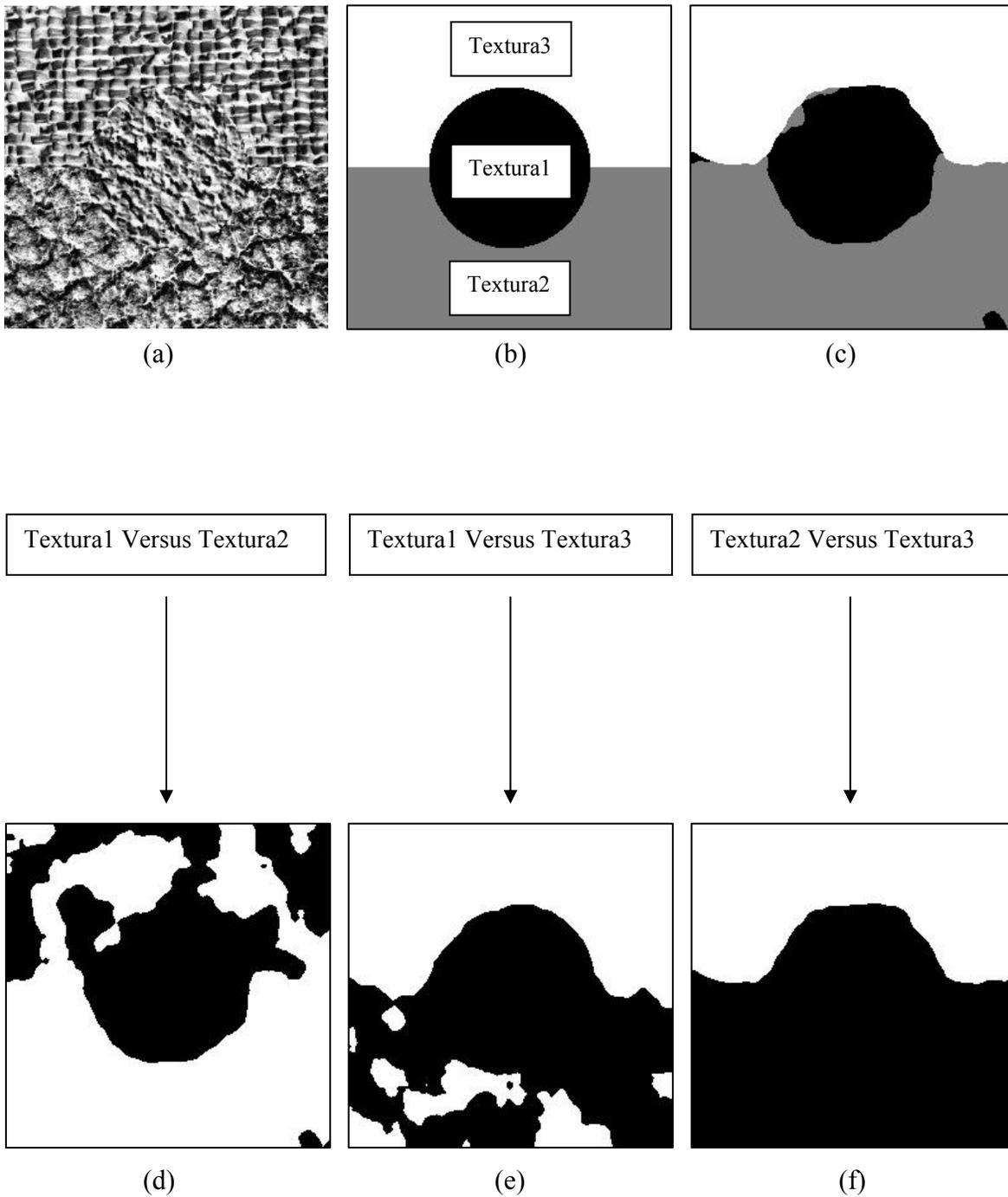


Figura 5.7: (a): mosaico de três texturas, (b): imagem referência, (c): resultado da segmentação. (d) (e) e (f): resultado da limiarização das bandas de máxima separabilidade conforme legendas.

5.6 Sumário

Nesse capítulo, foram relatadas as motivações e a fundamentação relacionada ao desenvolvimento do modelo experimental utilizado para discriminar pares de texturas. Apresentou-se o esquema visual associado ao modelo experimental, descreveu-se e justificou-se todas as etapas do modelo. Apresentou-se também o algoritmo combinatorial utilizado para discriminar conjuntos de mais que duas texturas além de ilustrá-lo com um exemplo prático.

CAPÍTULO 6

APLICAÇÕES DO MODELO EXPERIMENTAL

6.1 Introdução

Nesse capítulo, aplica-se o modelo experimental desenvolvido no capítulo 5 em um conjunto de mosaicos de texturas formados por duas e por cinco texturas extraídas de Brodatz [3], de MIT Vision Texture Database [89] e de MeasTex Image Texture Database [90]. Os resultados obtidos são comparados com vários métodos existentes e apresentados nas tabelas comparativas 6.1 e 6.2. A partir dos percentuais mostrados nas tabelas comparativas, pode-se observar que, no caso de pares de texturas, o método aqui desenvolvido mostra-se em média superior a todos os outros, enquanto que em mosaicos com cinco e com dez texturas sua superioridade foi bastante expressiva.

A maior parte das texturas utilizadas na comparação de modelos de discriminação de texturas foram extraídas do álbum “*A Photographic Album for Artists and Designers*” de Phil Brodatz [3], lançado em 1966 e consistindo de uma coletânea de 112 fotos de texturas naturais, tais como casca de árvores, pedras, grama, entre outras, e algumas texturas manufaturadas pelo homem, como por exemplo tecidos, telas de aço, etc.

As imagens apresentadas nesse álbum tornaram-se, com o passar do tempo, uma espécie de padrão de comparação de algoritmos. Pesquisadores do mundo todo têm utilizado essas texturas como referência, pois representam um conjunto de texturas de fácil acesso, razoavelmente grande e de variados tipos, para validar seus métodos.

Não fugindo a regra, nesse capítulo opta-se por utilizar algumas dessas texturas a fim de avaliar o desempenho do método apresentado no capítulo 5 e compará-lo a resultados encontrados na literatura.

A apresentação dos resultados está dividida em duas partes. Na seção 6.3 estão apresentados os resultados relacionados à segmentação de mosaicos de pares de texturas e uma discussão dos resultados, enquanto que na seção 6.4, estão apresentados os resultados da segmentação feita em mosaicos formados por cinco texturas e por dez texturas e também a discussão dos resultados. Na seção 6.2, apresenta-se a descrição de alguns trabalhos que tratam a questão da comparação da performance de métodos de segmentação de texturas encontrados na literatura.

6.2 Trabalhos Anteriores

Provavelmente o estudo mais completo publicado até hoje, comparando abordagens baseadas em filtragem, foi feito no artigo de Randen e Husoy [71], de 1999. Antes desse trabalho, haviam poucas comparações entre esquemas de extração de atributos de textura. Weszka et al. [86] compararam espectro de potência de Fourier, estatísticas de segunda ordem dos tons de cinza, estatísticas da matriz de co-ocorrência e “*gray-level run length statistics*”. Eles concluíram que os atributos oriundos de co-ocorrência eram os melhores entre todos esses. Num estudo feito por Connors e Harlow [15], co-ocorrência foi o melhor quando comparado com “*run length difference*”, “*gray-level difference density*” e o espectro de potência.

Du Buf et al. [22], relataram que vários atributos têm praticamente o mesmo desempenho. Eles avaliaram atributos oriundos de co-ocorrência, dimensão fractal, “*transform and filter bank features*”, frequências locais dominantes e orientação, número de extremos locais por unidade de área e integração curvilínea.

Ojala et al [59], compararam “*gray-level difference*”, filtros de Laws [41], covariância, padrões binários locais e pares de atributos complementares. Eles concluíram que “*gray-level difference*” tem o melhor desempenho enquanto que os atributos oriundos dos filtros de Laws obtiveram um desempenho muito pobre.

Atributos oriundos de filtragem têm sido comparados a atributos oriundos de co-ocorrência em muitos estudos, com diferentes conclusões. Strand e Tax [76] concluíram que atributos de co-ocorrência tinham melhor desempenho, enquanto que Laws [41], Pietikäinen et al. [64], e Claussi e Jernigan [14] obtiveram conclusão contrária. Esses resultados conflitantes devem-se, provavelmente, ao uso de diferentes modelos experimentais, diferentes imagens teste e diferentes esquemas de filtragem.

No artigo de Randen e Husoy [71], encontra-se uma extensa avaliação de abordagens relacionadas a atributos oriundos de filtragem. Nesse trabalho, os autores procuraram garantir certa padronização nos procedimentos, resguardadas suas particularidades, e os aplicaram num conjunto de imagens do tipo Brodatz, MIT Vision Texture Database [89] e MeasTex Image Texture Database [90] a fim de obter comparabilidade dos métodos.

Devido a enorme quantidade de modelos de discriminação de texturas utilizadas nas comparações apresentadas no trabalho de Randen e Husoy [71], utiliza-se aqui, como referência, os resultados e as imagens testes desse trabalho, que por sinal encontram-se disponibilizadas na Internet, no site, <http://www.ux.his.no/~tranden/>, a fim de avaliar e comparar o desempenho do modelo experimental aqui proposto. Isso evita o uso de imagens capturadas por equipamentos diferentes e sob condições diversas, propiciando uma padronização dos testes e comparabilidade dos resultados com os demais métodos.

Em Pagamisse e Dutra [61], foi utilizado um esquema experimental similar ao desse trabalho utilizando a transformada wavelet diádica, mas num modelo matemático com apenas duas direções, horizontal e vertical. Os resultados obtidos e uma análise dos mesmos encontra-se nas tabelas comparativas e nas conclusões que se seguem após os experimentos.

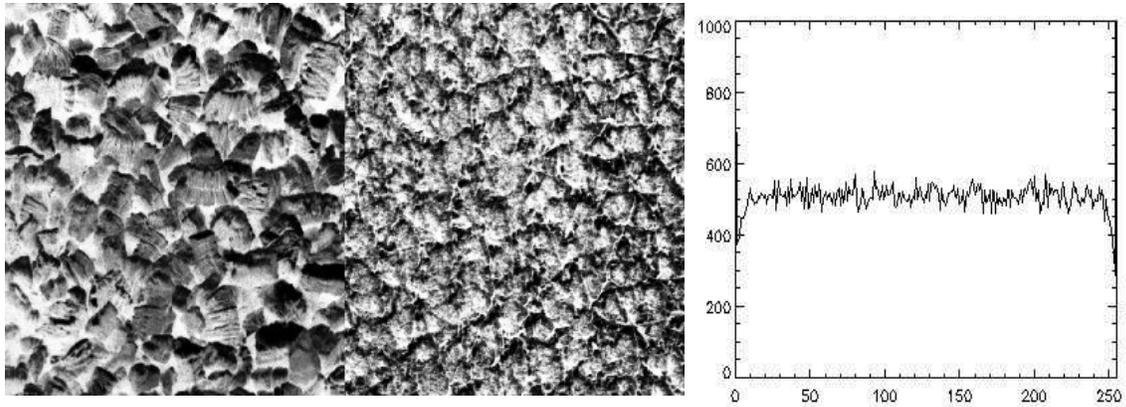
6.3 Segmentação de Pares de Texturas

Conforme mencionado no capítulo 5, obtendo-se bons resultados na segmentação de pares de texturas, então será mais fácil reproduzir esses bons resultados quando se deparar com uma imagem que apresente uma quantidade maior de texturas. Os pares de texturas da figura 6.1 são utilizados em Randen e Husoy [71] em sua extensa avaliação e os resultados obtidos pelos inúmeros métodos por eles avaliados, encontram-se apresentados na tabela 6.1.

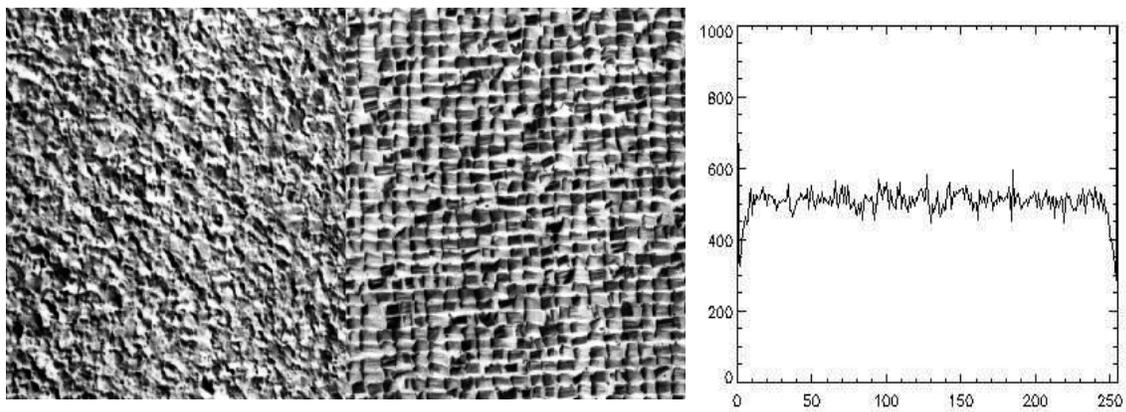
A fim de que as imagens da figura 6.1 não sejam discrimináveis com simples estatísticas, elas foram “homogeneizadas” com relação aos quatro primeiros momentos, isto é, todas possuem a mesma média, a mesma variância, a mesma kurtose e a mesma skewness. Como resultado dessa homogeneização, pode-se observar como os histogramas globais na figura 6.1 são praticamente idênticos e unimodais.

Na figura 6.2, apresenta-se as bandas de máxima separabilidade juntamente com seus respectivos histogramas. Note que nesse momento os histogramas são bimodais, ressaltando a possibilidade de segmentação dos pares de texturas com uma simples limiarização.

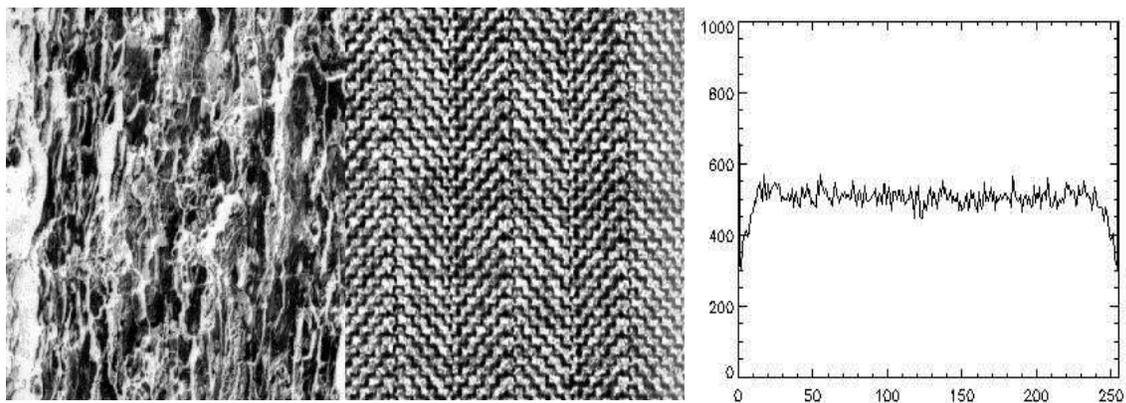
Os passos utilizados para a geração das bandas de máxima separabilidade são aqueles descritos no capítulo 5 quando da descrição do modelo experimental. As amostras das texturas são de tamanho 32x32 pixels, extraídas diretamente das



a

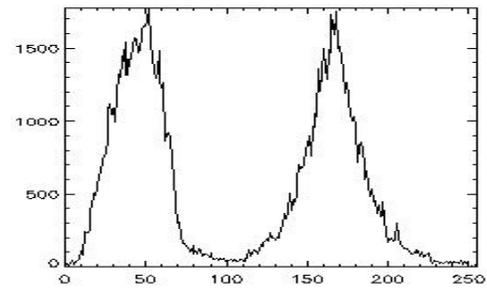
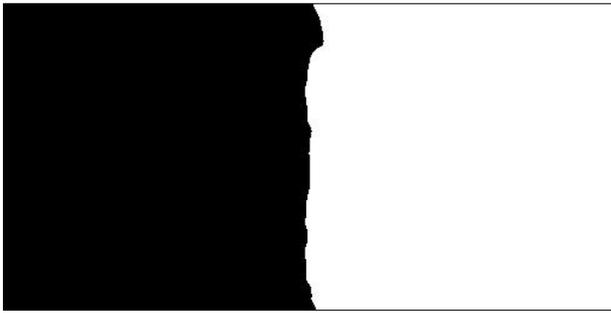
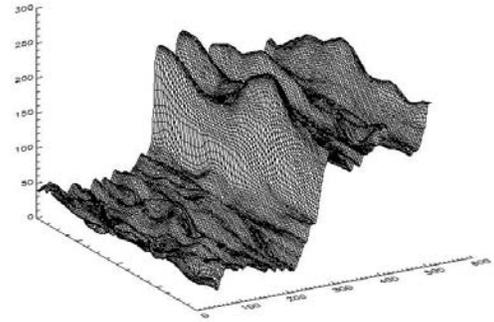


b

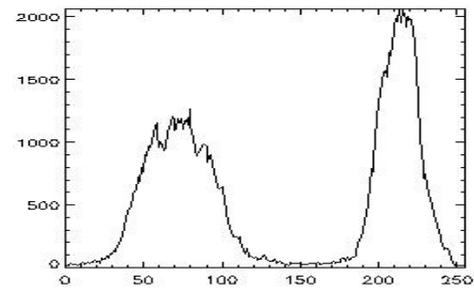
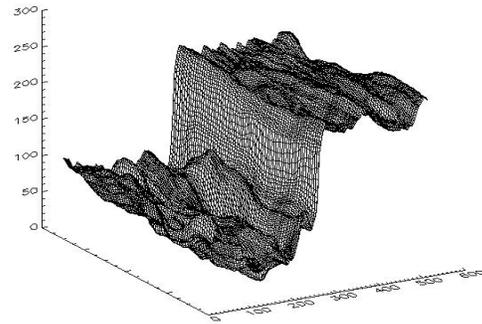
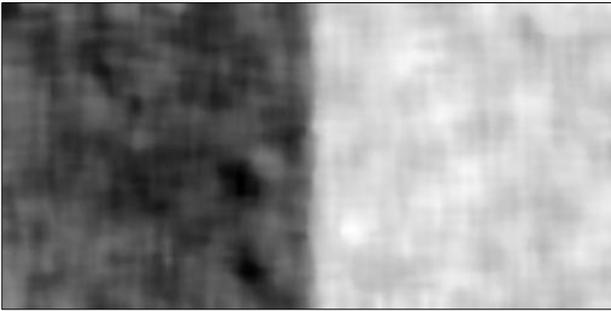


c

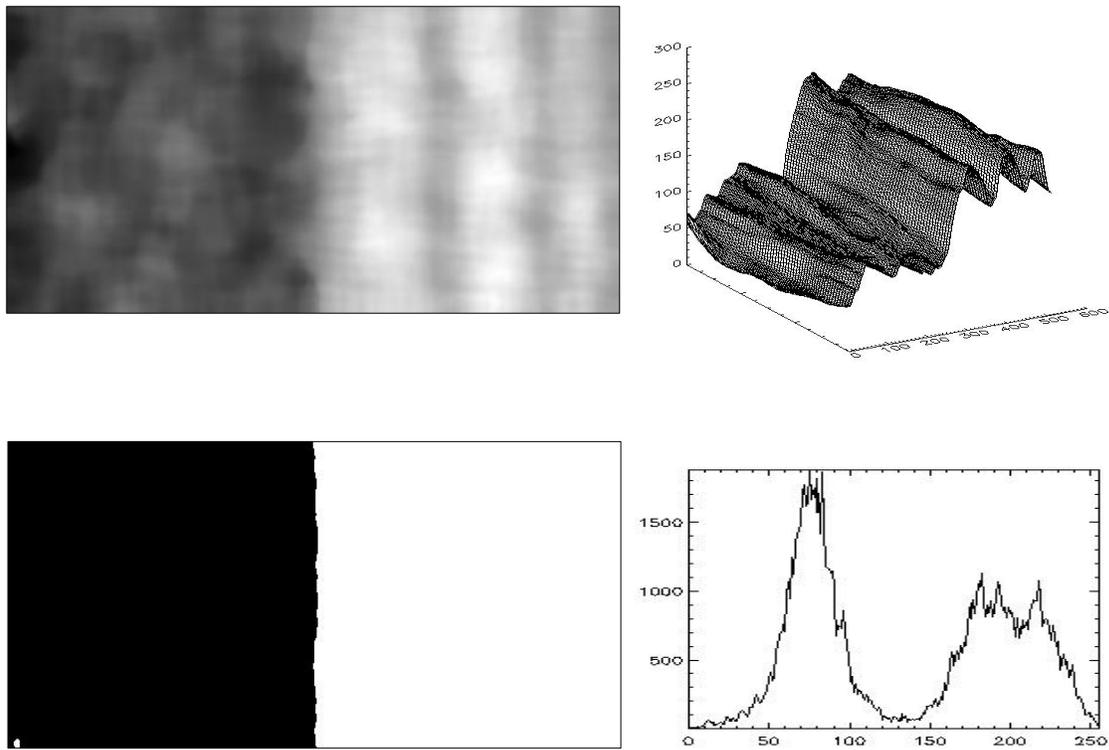
Figura 6.1. (a) par de texturas D05D92 e histograma, (b) par de texturas D04D84 e histograma e (c) par de texturas D12D17 e histograma.



(a)



(b)



(c)

Figura 6.2: No sentido horário, banda de máxima separabilidade, superfície dos tons de cinza, threshold e histograma da banda de máxima separabilidade dos pares de texturas da figura 6.1.

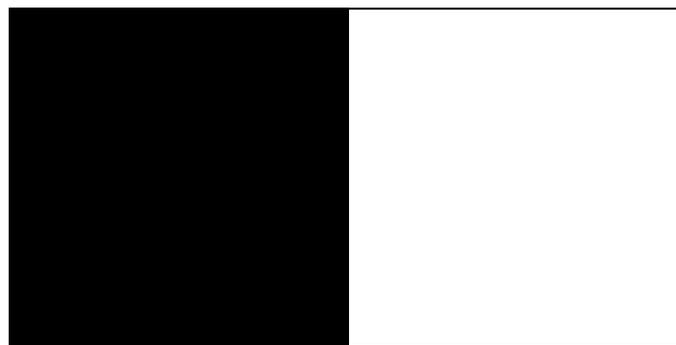


Figura 6.3: Imagem referência, 512x256 pixels, para pares de texturas.

imagens com o uso do mouse para marcar o ponto inicial (x,y) da amostra. As amostras devem refletir da melhor maneira possível o comportamento “textural” da imagem, desse modo, foram tomadas de modo a serem o mais representativamente possível.

No caso dos pares apresentados na figura 6.1, constata-se que as texturas apresentam aspectos estruturais que as diferenciam, tais como, diferenças de escala, de orientação e repetitividade de texel gerador mais presente em algumas do que em outras. As texturas componentes de cada par são de tamanho 256x256 pixels, em 8 bits, extraídas de Brodatz e formadas por:

- figura 6.1 a: formada pelas texturas D5 e D92
- figura 6.1 b: formada pelas texturas D4 e D84
- figura 6.1 c: formada pelas texturas D12 e D17

As texturas em 6.1 a-) se caracterizam principalmente por diferenças na escala, enquanto que as texturas em 6.1 b-) pela repetitividade e nas texturas em 6.1 c-) pelas diferenças em orientação, escala e repetitividade.

6.3.1 Discussão dos Resultados

Apesar do método aqui apresentado não ser o de melhor desempenho quando aplicado à textura D12D17, a diferença, nesse caso, é muito pequena, além disso, valores próximos e abaixo de 1%, estão num limite muito sutil de erro. Quanto ao bom desempenho de JMS, JU e JF, que são filtros otimizados para o caso de duas texturas [71], no par D04D84, deve-se ao fato que esse par de texturas tem um alto grau de repetitividade na sua formação, isto é, ambas possuem um “texel” gerador bastante homogêneo. Como esses filtros são construídos de modo a otimizar (prever) a saída de energia, quanto mais homogênea for a textura, mais eficiente será o resultado. Contudo nem sempre essa é a situação encontrada na prática, podendo-se notar os resultados não tão exatos nos demais pares de texturas.

Método	D04D84	D05D92	D12D17	Média
Filtros de Laws	24,8	8,6	3,3	12,23
Filtros Anel/Wedge	3,0	18,4	4,5	8,63
Banco de Filtros Diádicos de Gabor	3,6	11,3	2,0	5,63
DCT	6,4	2,5	2,2	3,7
Wavelet Frame	4,4	4,7	0,4	3,17
QMF	6,7	6,0	0,8	4,5
Co-ocorrência	1,9	3,3	4,8	3,33
AR	11,3	3,0	1,9	5,4
Gabor full rate	6,5	15,6	1,2	7,77
Daubechies 4 full rate	5,7	8,2	0,6	4,83
Daubechies 6 full rate	4,7	6,1	0,6	3,8
Daubechies 8 full rate	4,9	4,7	1,3	3,63
Daubechies 10 full rate	4,4	5,1	1,2	3,57
F 16 b, full rate	8,1	8,2	0,8	5,7
F_2_1_smpl, full rate	7,2	7,4	2,3	5,63
F_2_1_09, full rate	6,7	6,0	2,1	4,93
Mahalanobis-Singh	1,3	1,4	24,6	9,1
Fisher, closed form	1,4	7,4	4,6	4,46
Fisher, gradient	1,9	1,6	3,0	2,2
Dunn	2,8	1,0	25,4	9,7
Eigenfilter	4,0	4,7	3,7	4,1
Opt. Repr. Filter Gabor	5,5	36,2	5,8	15,8
JMS	0,7	28,6	8,3	12,5
JU	0,7	5,1	2,1	2,6
JF	0,7	5,1	2,1	2,6
Optimal Gabor filter (best)	5,3	10,8	3,3	6,5
Prediction error filter	0,9	31,7	4,6	12,4
BackProp. NN mask size=11	32,1	43,6	30,0	35,2
BackProp. NN mask size=21	30,8	48,6	28,8	36,1
4,6,10 filter opt. Gabor filter bank (melhor)	3,8	12,8	3,8	6,8
Diádicas 2 direções	0,8	2,6	2,4	1,9
Diádicas 4 direções	0,6	1,1	1,1	0,9

Tabela 6.1: Porcentagem de erro para diversos métodos de segmentação aplicados aos pares de texturas D04D84, D05D92, D12D17. Resultados extraídos de Randen e Husoy [78] e Pagamisse e Dutra [37].

Na maioria das vezes, o percentual de erro obtido na segmentação é proveniente da região situada na borda entre as texturas. Nessa região existe uma área de “confusão”, isto é, uma área em que a intensidade da energia resultante se encontra entre as intensidades das energias das duas texturas. Isso pode ser verificado nas superfícies que aparecem na figura 6.2. Apenas para efeito de comparação, sabendo que uma única coluna num mosaico de duas texturas, representa 256 pixels, isto é, 0,39% do total de pixels, então, duas colunas e meia, mal classificadas, na borda entre as duas texturas, irão representar aproximadamente 1% de erro.

Pode-se concluir, pelo menos nesses três pares de texturas que o método aqui apresentado é muito superior à maioria dos outros métodos da tabela 6.1 e em média melhor do que todos. Uma de suas qualidades está no fato que os valores obtidos não diferem muito, ou seja, ele obtém bons resultados em todos os mosaicos, mostrando ser mais robusto que os demais.

Cabe observar o menor percentual de erro do modelo utilizando a transformada wavelet discreta com quatro direções, comparado ao modelo utilizando apenas duas direções. Isso mostra que vale a pena aumentar a complexidade do modelo e também a complexidade computacional. Pode-se perguntar se a inclusão de mais direções implicaria em melhores resultados. A experiência mostrou que praticamente não há ganhos, não compensando o esforço computacional necessário. Além disso, existe a chamada “maldição da dimensionalidade”, que diz que a partir de certo momento, adicionar mais informação (bandas) pode prejudicar a segmentação ao invés de melhorá-la.

6.4 Segmentação de Mosaicos de Cinco Texturas e de Dez Texturas

Seguindo a apresentação de Randen e Husoy, com relação ao tipo de testes efetuados, aplicou-se o modelo experimental aqui desenvolvido para

segmentar mosaicos com cinco texturas e mosaicos com dez texturas. Em problemas reais, dificilmente depara-se com uma grande quantidade de texturas que necessitem ser diferenciadas. Esses mosaicos foram obtidos da mesma forma que os pares de texturas utilizados inicialmente, isto é, do site (<http://www.ux.his.no/~tranden/>). A figura 6.4 apresenta os mosaicos de cinco texturas, enquanto que as figuras 6.6 e 6.7 apresentam os mosaicos de 10 texturas, o resultado da segmentação e a imagem referência do mosaico. A tabela 6.2 mostra a performance de vários métodos utilizados para segmentá-los, conforme Randen e Husoy, juntamente com os resultados aqui obtidos. As texturas desses mosaicos não são oriundas exclusivamente de Brodatz, mas também do MIT Vision Texture Database [89] e do MeasTex Image Texture Database [90].

As amostras usadas para se obter os parâmetros para a geração das bandas ótimas são de tamanho 32x32. Elas foram obtidas diretamente das imagens teste, usando-se o cursor associado ao mouse, simplesmente clicando sobre a imagem e utilizando essa posição como a posição inicial da amostra, tendo-se a preocupação de que essas amostras fossem representativas das texturas.

Os mosaicos 5c, 5v, 5v2, 5v3, 5m (256x256), Nat10 e Nat10v (256x640), conforme designação de Randen e Husoy, foram formados pelas seguintes texturas:

- Mosaico 5c: D77, D84, D55, D53 e D24, [3].
- Mosaico 5v: Fabric.000, Fabric.017, Flowers.002, Leaves.006 e Leaves.013, [Mit].
- Mosaico 5v2: Fabric.009, Fabric.016, Fabric.019, Flowers.005 e Food.005 [Mit]
- Mosaico 5v3: Fabric.007, Fabric.009, Leaves.003, Misc.002 e Sand.000 [Mit].
- Mosaico 5m: Asphalt.000, Concrete.001, Grass.002, Misc.002 e Rock.005, [Meas].
- Mosaico Nat10: D04, D09, D19, D21, D24, D28, D29, D36, D37, D38, [3].
- Mosaico Nat10v: Fabric.009, Fabric.016, Fabric.019, Flowers.005, Food.005, Leaves.003, Misc.000, Misc.002, Sand.000, Stone.004, [Mit] [Meas].

Método	5 c	5 v	5 v2	5 v3	5 m	Média 5 text.	Nat10	Nat10 v
Filtros de Law	9,7	25,7	32,4	27,3	25,7	24,1	41,9	37,8
Filtros Anel/Wedge	14,6	35,5	28,9	35,5	22,4	27,3	44,5	48,3
Banco de Filtros Diádicos de Gabor	10,7	34,8	22,6	25,2	24,6	23,6	32,3	47,9
Banco de filtros de Gabor	8,2	34,0	25,8	36,9	28,4	26,7	39,7	54,8
DCT	13,2	27,0	25,5	37,8	22,6	25,2	38,2	33,0
Co-ocorrência	9,9	27,0	26,1	51,1	35,7	30,0	35,3	49,1
AR	19,6	19,4	23,0	23,9	24,0	22,0	56,7	28,7
Daubechies 4	8,7	22,8	25,0	23,4	21,8	20,3	40,9	30,1
F 16 b	8,7	18,9	23,3	18,4	17,2	17,3	39,8	28,5
Eigenfilter	12,1	27,0	25,0	29,7	23,4	23,4	37,1	33,3
Opt. Repr. Gabor filter bk.	7,2	60,5	36,4	27,4	42,2	34,7	41,9	67,5
Prediction error filter	14,7	31,1	27,0	34,0	42,5	29,9	44,6	36,7
4-filter opt. Gabor filter bk.	24,7	54,5	34,2	33,3	34,0	36,1	52,8	51,4
6-filter opt. Gabor filter bk.	16,3	34,5	33,9	33,6	29,3	29,5	50,3	45,4
10-filter opt. Gabor filt. bk.	16,1	29,4	31,3	34,5	27,8	27,8	39,9	37,3
JMS	16,9	36,3	32,7	41,1	43,0	34,0	59,7	49,9
JU	12,7	33,0	26,5	34,3	43,4	30,0	35,9	30,5
BackProp. NN mask size=11	47,4	65,0	71,4	67,2	55,7	61,2	69,3	70,2
BackProp. NN mask size=21	72,4	68,7	69,0	74,8	70,0	71,0	76,5	84,3
Diádicas 2 direções	11,7	17,7	22,3	27,8	18,2	19,5	41,5	43,4
Diádicas 4 direções	6,9	9,5	15,7	17,9	10,2	12,0	17,7	27,1

Tabela 6.2: Porcentagem de erro para diversos métodos de segmentação aplicados aos mosaicos 5c, 5v, 5v3, 5m, Nat10 e Nat10v. Resultados extraídos de Randen e Husoy [71] e Pagamisse e Dutra [61].

6.4.1 Discussão dos Resultados

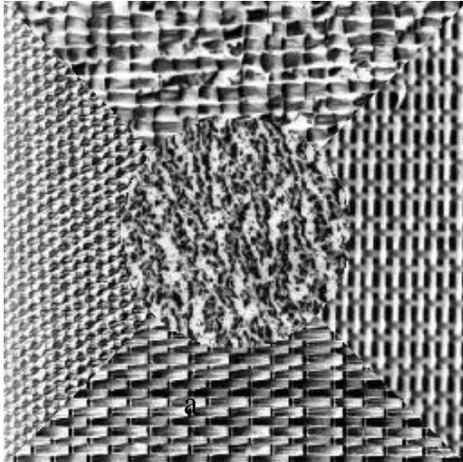
A partir dos resultados apresentados na tabela 6.2, pode-se observar o bom desempenho do método proposto quando comparado aos demais métodos. Cabe dizer que os resultados que foram apresentados são os melhores entre algumas tentativas de escolha do posicionamento adequado das amostras usadas para

treinamento. Esse procedimento é necessário em texturas irregulares. O uso de amostras maiores diminuiria esse problema. Tal procedimento é o mesmo que foi utilizado por Handen e Husoy. A figura 6.5 apresenta o resultado da segmentação dos mosaicos com cinco texturas assim como uma imagem erro correspondente a cada mosaico segmentado. As figuras 6.6 e 6.7 apresentam os mosaicos de dez texturas, o resultado da segmentação, assim como a imagem referência.

Em alguns mosaicos, tais como o mosaico 5c, os experimentos mostraram que a variabilidade nos resultados é pequena, não havendo muita dependência das amostras utilizadas. Contudo, nos demais mosaicos, encontra-se uma variabilidade bem maior. Isso se deve a natureza homogênea das texturas em 5c e a não homogeneidade observada em algumas texturas nos demais mosaicos. Esse comportamento influi não só na variabilidade, mas também no percentual de erro. Essa alta variabilidade pode ser também se observada nos demais métodos testados por Randen e Husoy, fazendo que o desempenho decaia de modo diretamente relacionado.

Com relação aos mosaicos de dez texturas, o método mostrou excelentes resultados, principalmente no mosaico Nat10, onde superou os demais com larga margem. Da mesma forma que nos mosaicos de cinco texturas, isso se deve à maior homogeneidade das texturas do mosaico Nat10. Pode-se notar que no mosaico Nat10v, apesar do método apresentar melhor rendimento que os demais, sua superioridade não foi tão expressiva. Isso se deve à questão da não homogeneidade das texturas componentes desse mosaico e inclusive à semelhança visual entre algumas delas.

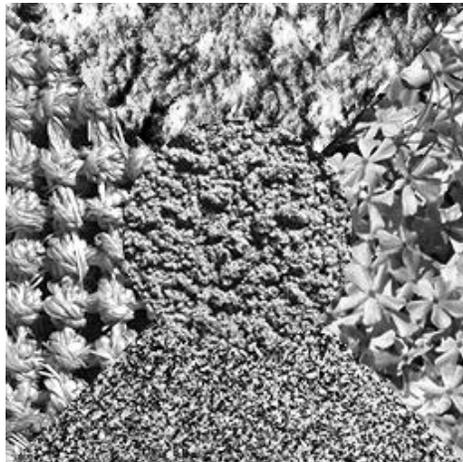
Cabe observar também, mais acentuadamente nos mosaicos com cinco texturas e com dez texturas que nos mosaicos de duas texturas, o bom rendimento do método com quatro direções, comparado ao método utilizando a transformada wavelet diádica com duas orientações. Apesar da não homogeneidade de algumas de texturas, o método com quatro orientações é bem mais robusto, sustentando seus índices de acerto enquanto aumenta-se o número de texturas a serem segmentadas, o que não ocorre com o método com duas direções.



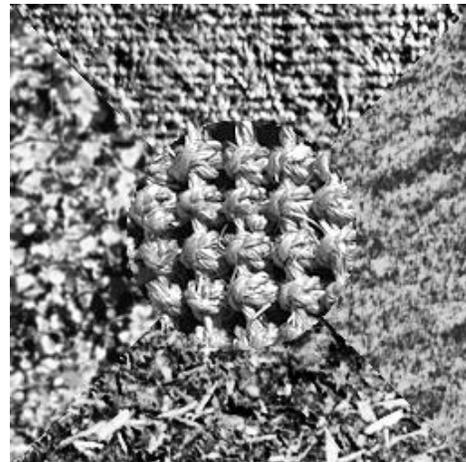
a



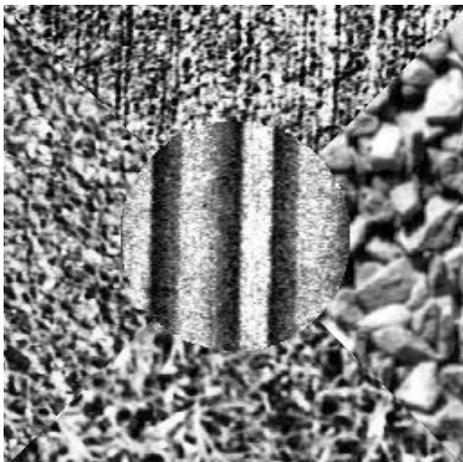
b



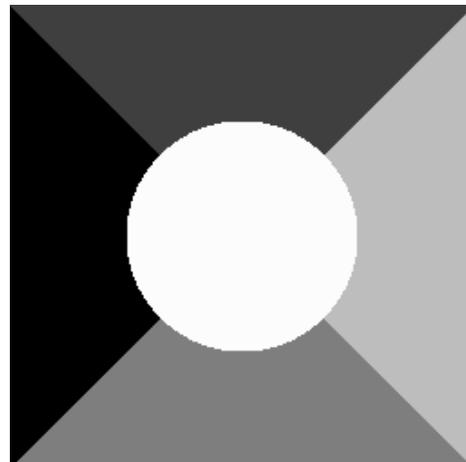
c



d



e



f

Figura 6.4: Mosaicos de 5 texturas, a-) 5c, b-) 5v, c-) 5v2, d-) 5v3, e-) 5m, f-) Imagem referência.

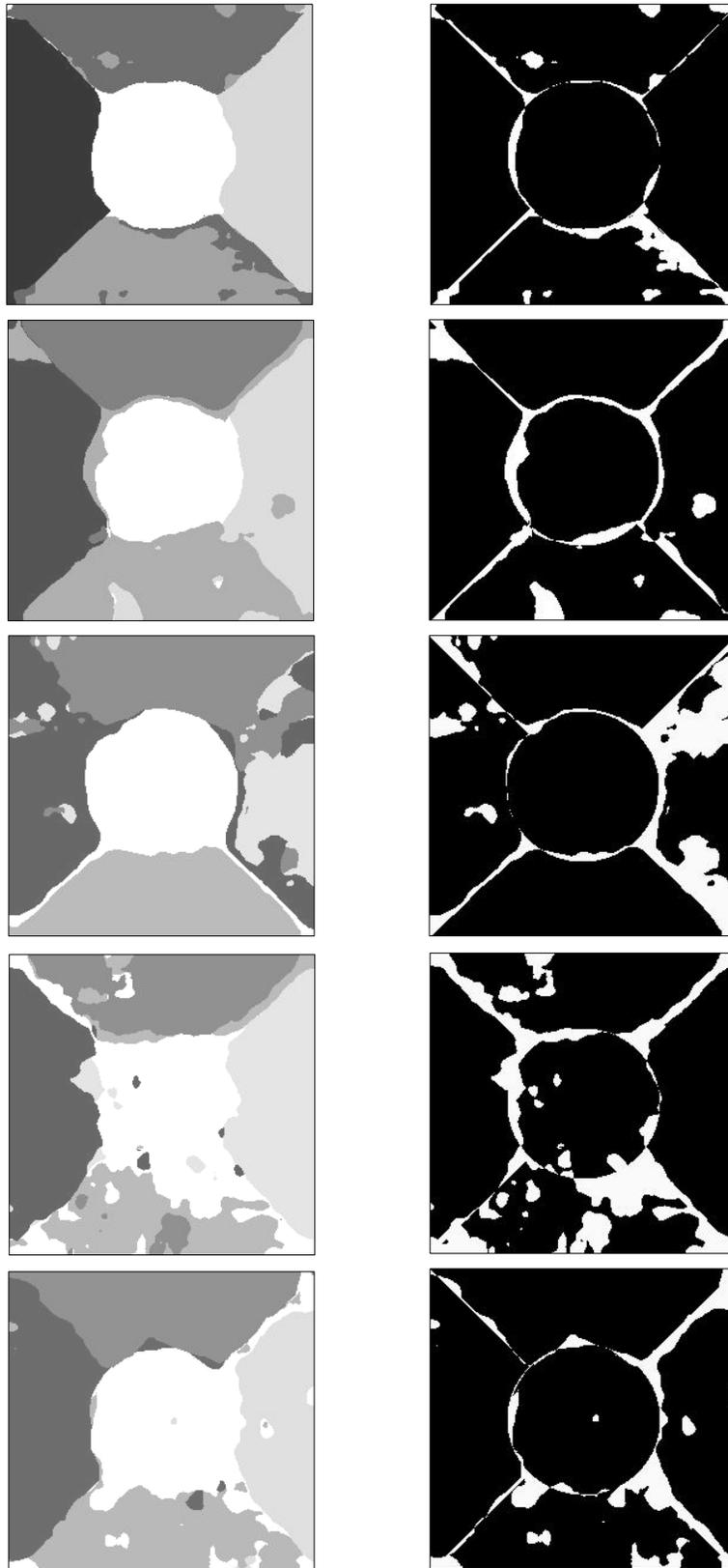


Figura 6.5. Resultado da segmentação dos mosaicos 5c, 5v, 5v2, 5v3, 5m (de cima para baixo), e respectivas imagens erro.

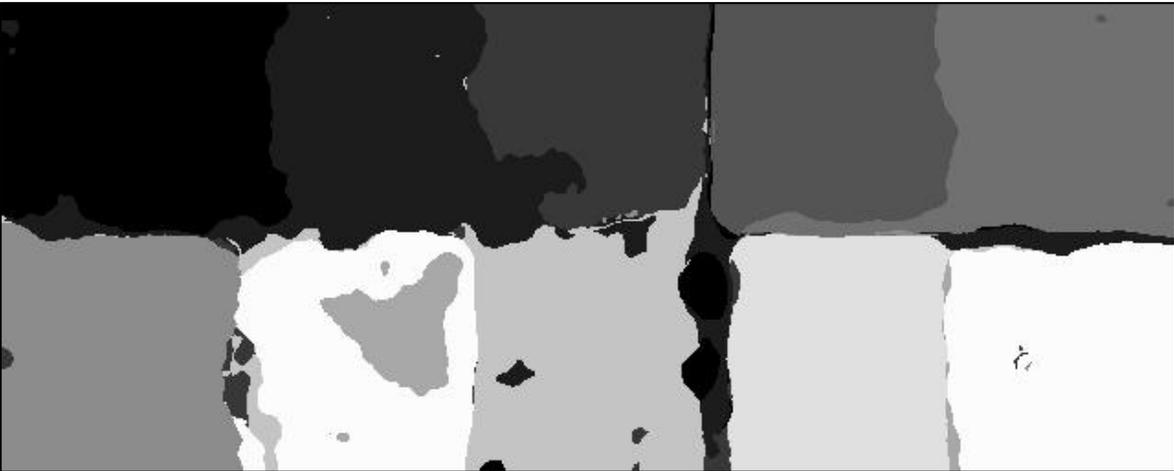
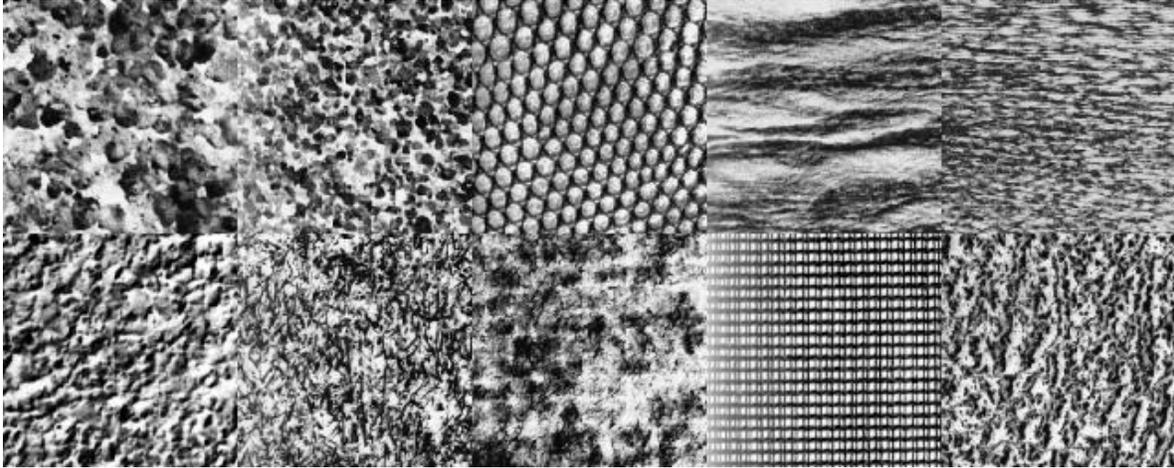


Figura 6.6: Superior: Mosaico de 10 texturas, Nat10; intermediária: Imagem segmentada; inferior: Imagem referência.

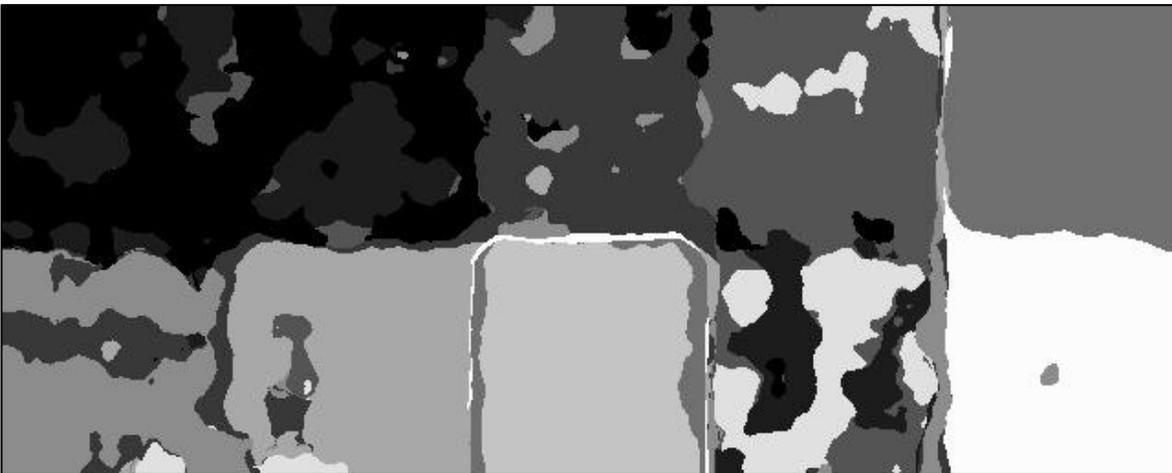
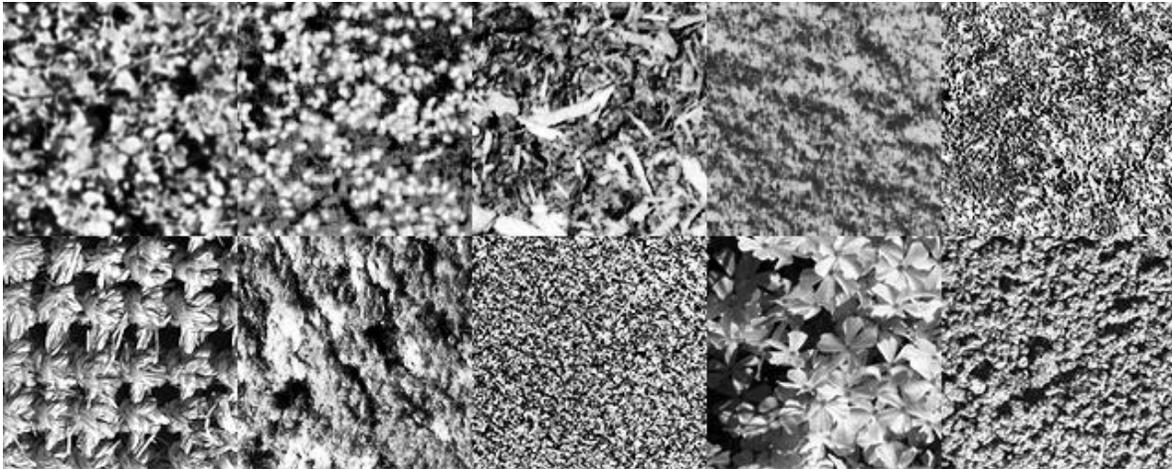


Figura 6.7: superior: Mosaico Nat10v; inferior: Resultado da segmentação.

6.5 Uma Aplicação em Spring

O objetivo desse trabalho foi o de desenvolver uma técnica de discriminação de texturas para ser utilizada em Processamento Digital de Imagens. Em princípio, não foi preocupação aqui dar tanta ênfase a técnicas de classificação, mas sim desenvolver um modelo matemático de representação utilizando a transformada wavelet e mostrar que esse modelo é adequado à tarefa de discriminação de texturas, apesar da apresentação de um algoritmo para realizar essa tarefa, o qual obteve bons resultados. Esses bons resultados se

devem à qualidade das bandas geradas pelo método proposto. Desse modo, ao invés de utilizar o algoritmo combinatorial, pode-se simplesmente gerar os atributos (bandas) para que o usuário, em posse dessas bandas, utilize um classificador de sua preferência.

A fim de testar o que foi proposto no parágrafo acima, tomou-se o mosaico Nat-5b, formado por cinco texturas e obtido da *homepage* de T. Randen, e gerou-se as bandas de máxima separabilidade para o nível 1, dez bandas nesse caso, conforme a descrição feita no capítulo 5. Esse mosaico foi segmentado utilizando o software Spring.

O SPRING é um Sistema de Informações Geográficas (GIS), no estado-da-arte, com funções de processamento de imagens, análise espacial, modelagem numérica de terreno e consulta a bancos de dados espaciais, desenvolvido no INPE, Instituto Nacional de Pesquisas Espaciais e é distribuído livremente.

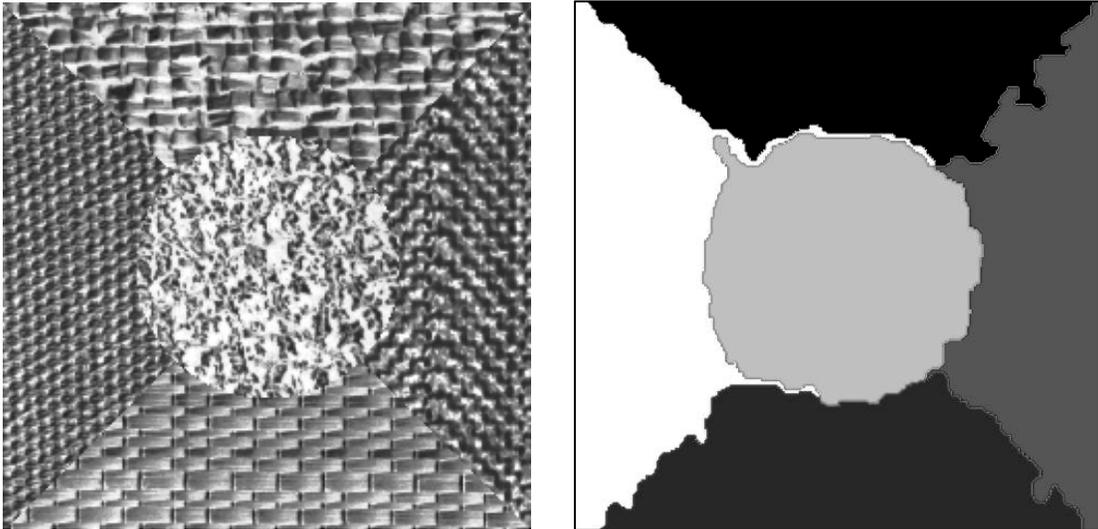


Figura 6.8: Esquerda: Mosaico Nat-5b; Direita: resultado da segmentação usando o software Spring.

A segmentação foi feita utilizando crescimento de regiões com grau de similaridade 2 e 9000 pixels como limite mínimo por região, supervisionada utilizando mínima distância de Bhattacharyya [23]. Pode-se verificar pela figura 6.8

o bom resultado da segmentação obtida usando o Spring, comprovando a qualidade do método aqui desenvolvido.

6.6 Sumário

Nesse capítulo, aplicou-se o método desenvolvido em mosaicos de texturas obtidas de Brodatz, de [89] e de [90], formados por duas, cinco e dez texturas. Comparou-se os resultados obtidos com os inúmeros resultados apresentados em Randen e Husoy [71]. A relevância de fazer a comparação com esse trabalho deve-se ao fato que, muito provavelmente, ele seja um dos mais atuais e completos que existem na literatura. As imagens utilizadas aqui foram as mesmas imagens utilizadas por Randen e Husoy, as quais se encontram disponibilizadas na Internet e podem ser obtidas por download.

Os resultados obtidos mostraram-se superiores a todos os demais métodos, com margens de erro inferiores ou pelo menos em média inferiores, mostrando a qualidade e relevância do método.

Apresentou-se também o resultado de uma segmentação realizada utilizando as bandas de máxima separabilidade geradas pelo método. Essa segmentação foi realizada utilizando o software Spring e apresentou também excelentes resultados, mostrando que uma das qualidades do método é a de gerar atributos de textura de boa qualidade para o usuário final.

Capítulo 7

Conclusões e Trabalhos Futuros

7.1 Conclusões

Neste trabalho, propomos desenvolver uma técnica de segmentação de texturas usando a transformada Wavelet. A análise de texturas, continua sendo um tópico de grande interesse, haja vista a enorme quantidade de trabalhos sendo publicados sobre o assunto. Mesmo assim, ao se deparar com problemas práticos, nem sempre as teorias desenvolvidas conseguem obter bons resultados. Contudo, cada novo estudo acrescenta algum novo conhecimento que avança a discussão do problema.

A transformada Wavelet, apesar de bem recente, tornou-se praticamente uma panacéia. Inúmeros pesquisadores de inúmeras áreas estão nesse momento buscando alguma forma de utilizá-la. Em muitos casos, velhas teorias estão sendo reescritas com uma formalização usando wavelets. Nesse trabalho, isso também não é diferente pois a formulação para as wavelets que usamos, nasceu das idéias de Mallat, influenciado pelas teorias de detecção de bordas de Canny e os cruzamentos por zero de Marr-Hildreth, comuns na engenharia elétrica.

O diferencial desse trabalho, é que essa implementação das wavelets diádicas, por não realizar a decimação no sinal, presta-se de modo mais adequado ao reconhecimento de padrões, pois, preserva a localização dos pixels,

é invariante por translações e possui boas propriedades matemáticas, além é claro das propriedades tradicionais das wavelets que as tornam tão desejáveis.

A qualidade desse trabalho é a de acreditar que a decomposição wavelet (representação) dos sinais dada por essa formulação, adicionada por duas direções, seria adequada à tarefa de segmentar texturas e, para provar essa afirmação, desenvolvemos um modelo experimental que, juntamente com a decomposição wavelet, introduz outras operações cuja motivação advém das propriedades intrínsecas das texturas. Esse modelo experimental mostrou excelente desempenho quando comparado a outros métodos. Nesse ponto, não podemos deixar de elogiar o extenso estudo realizado por Randen e Husoy, sem o qual esse e mais alguns outros trabalhos não teriam em que se espelhar; pelo menos não de modo tão abrangente. Além disso, sua iniciativa de disponibilizar os mosaicos de texturas para download, está possibilitando a muitos pesquisadores poderem validar seus métodos, sem a preocupação de se saber como as imagens foram escaneadas, e outros detalhes dessa natureza que poderiam interferir na comparabilidade dos resultados. Não sabemos qual a motivação deles ao escolher tais mosaicos, contudo é o que temos de melhor até o momento.

7.2 Trabalhos Futuros

Voltando ao método aqui desenvolvido, pode-se dizer que ele passou no teste das texturas do tipo Brodatz proposto por Randen e Husoy, contudo, existe a necessidade de aplicá-lo em problemas reais para ver como se comporta e quais adaptações serão necessárias, afinal, texturas são consideravelmente dependentes da aplicação. Podemos elencar essa tarefa como uma pesquisa futura.

Outra pesquisa a ser feita, utilizando a capacidade de gerar atributos pela transformada wavelet diádica, seria a de desenvolver alguma forma não

supervisionada de segmentação, já que o modelo aqui desenvolvido é essencialmente supervisionado.

Existe também muitas outras possibilidades e testes passíveis de serem feitas com o modelo experimental. Pode-se buscar outras famílias de wavelets, pode-se utilizar análise discriminante entre níveis, pode-se buscar informações sobre as primitivas que compõem as texturas e introduzir esses conhecimentos à priori de algum modo no modelo. É muito interessante observar que para certos pares de texturas, uma única escala e uma única direção já possibilita a discriminação de um par de texturas, e essa, é uma parte muito pequena de toda a informação que as texturas carregam, se pensarmos a decomposição wavelet como uma representação. O que isso significa? Que há informação redundante? Sim, pelo menos para a tarefa de segmentação...

Com a combinação adequada dos atributos gerados pela WT, pode-se buscar projetar sistemas de classificação de texturas que sejam sensíveis ou não à orientação e escala e que sejam ajustáveis pelo usuário. Esse é um problema típico em Geomorfologia, que se interessa em discriminar texturas em escalas mas não em orientação, enquanto que no estudo de ondas do mar, o interesse é na direção e não na escala. Em geral, essas tarefas são feitas por interpretação visual, abrindo a possibilidade de criar sistemas que realizam essas interpretações de modo automático.

Outra abordagem interessante, relacionada a esse trabalho, e que vem surgindo na literatura, são os chamado filtros de Volterra, que são filtros não lineares e que, de certo modo, englobam as operações lineares e não lineares do modelo desenvolvido. Pode-se dizer que o modelo experimental desenvolvido é um caso particular dos filtros de Volterra. A busca por abordagens não lineares, é sugerida por Stephane Mallat, em seu excelente livro, "A Wavelet Tour of Signal Processing" [46], ainda na introdução do livro, propõe,

"Try it non-linear. Linearity has long predominated because of its apparent simplicity. We are used to slogans that often hide the limitations of "optimal" linear procedures such as Wiener filtering or Karhunen-Loève bases expansions. Yet,

remarkably simple non-linear algorithms like thresholdings in adapted bases can outperform "optimal" linear procedures. The complexity of transient signal processing motivates the elaboration of adaptative non-linear algorithms, while trying to keep the mathematics under control".

Assim, métodos não lineares apresentam-se como um caminho um natural. Muitas vezes, difícil de ser seguido, principalmente pelas dificuldades teóricas e computacionais. Uma explanação mais aprofundada sobre filtros não lineares e principalmente filtros de Volterra, é dada no apêndice A desse trabalho.

A fim de tratar a questão da discriminação de texturas utilizando os filtros de Volterra, sugerimos algumas abordagens a serem testadas:

1- Edge quadrático.

As bordas concentram a maior parte das informações relevantes de uma imagem. Assim, da mesma forma como se procedeu no caso das wavelets, é possível "imitar" o procedimento linear da decomposição wavelets e sua relação com bordas, tentando obter melhores resultados com o procedimento quadrático de Volterra.

2- Multirresolução.

Até o momento não se encontrou na bibliografia trabalho algum que realize filtros de Volterra num contexto de multirresolução. Essa é uma abordagem que pode ser tentada levando em conta o que já foi feito em termos de multirresolução e filtragem linear. Seu ineditismo pode tanto ser uma vantagem como uma dificuldade.

3- Filtros de Volterra Casados.

Conforme dito anteriormente, não-linearidades são comuns em modelos óticos, além disso, os filtros lineares não respondem bem em sinais não-Gaussianos. Portanto, uma modelagem quadrática, em princípio, poderia capturar de modo mais eficiente às peculiaridades de diferentes texturas. Nessa

modelagem, alguns aspectos como isotropia, simetria, preservação de energia, e rotações, entre outras, podem ser levados em conta pelo modelo ao construir os coeficientes do filtro quadrático.

4- Relação entre filtros de Volterra e Wavelets.

Em [57], Nowak e Baraniuk mostraram, para sinais 1-dimensionais, que uma NST (nonlinear signal transform) completa de n -ésima ordem é capaz de realizar qualquer filtro de Volterra de mesma ordem. Tal realização é baseada em uma decomposição em dois passos, consistindo de uma expansão ortogonal do sinal seguida por uma transformação polinomial dos coeficientes resultantes do sinal. Por diversas razões, uma expansão wavelet é a mais interessante. Pode-se buscar a generalização dessa realização para sinais bidimensionais e suas possíveis aplicações na discriminação de texturas.

7.3 Considerações Finais

Dedicamos uma certa atenção à questão dos filtros de Volterra, pois apesar das dificuldades teóricas e computacionais inerentes a essa teoria, os desafios e os ganhos advindos dessa abordagem parecem-nos recompensadores. Obviamente que as demais possibilidades não serão relegadas a um segundo plano ou esquecidas, ao contrário, são possibilidades imediatas e que pelo fato de se utilizarem das ferramentas aqui desenvolvidas, terão maior facilidade de serem testadas e modificadas.

7.4 Sumário

Nesse capítulo, apresentaram-se algumas possibilidades de trabalhos futuros utilizando o ferramental desenvolvido. Apresentou-se também a necessidade e qualidades dos filtros não lineares, dando especial ênfase na

abordagem não-linear baseada nos filtros de Volterra. Tais filtros englobam os procedimentos relativos às filtragens e à não-linearidade que são utilizadas em nosso modelo experimental. Por sua generalidade e abrangência vem a ser um tópico com grandes possibilidades em pesquisas futuras.

Bibliografia

- [1] Ahmed, N. and Rao, K. R., "Orthogonal Transforms for Digital Signal Processing", Springer, Berlin, 1975.
- [2] Bazaraa, M. S., Shetty, C. M., "Nonlinear Programming", John Wiley & Sons, New York, 1979.
- [3] Brodatz, P., "A Photographic Album for Artists & Designers", Dover Publications, Inc., New York, 1966.
- [4] Burt, P. J. and Adelson, E. H., "The Laplacian Pyramid as Compact Image Code", IEEE Trans. Commun. Vol. 31, pp. 532-540, 1983.
- [5] Campbel, F., Kulikowski, J., "Orientation Selectivity of the Human Visual System", J. Phisiol., vol. 197, pp. 437-441, 1966.
- [6] Campbell, F. W. and Robson, J. G., "Application of Fourier analysis to the visibility of gratings", J. Physiology, 197, pp. 551-566, 1968
- [7] Canny, J. "A Computational Approach to Edge Detection". IEEE Trans. Patt. Anal. Machine Intell., vol. 8, pp. 679-698, 1986.
- [8] Chellappa, R., Chatterjee, S., "Classification of Textures using Gaussian Markov Random Fields", IEEE Trans. Acoust. Speech Signal Processing, Vol. 33, pp 959-963, 1985.
- [9] Chellappa, R., Kashyap, R. L., Manjunath, B. S., "Model-based Texture Segmentation and Classification", Handbook of Pattern Recognition and Computer Vision, pp 277-310, World Scientific, Singapore, 1993.

- [10] Chen, J., Jain, A. K., "A Structural Approach to Identify Defects in Textured Images", Proc. IEEE Int. Conf. On Syst., Man, Cyb, Beijing, pp 29-32, 1988.
- [11] Chiang, H. H., Nikias, C. L. and Venetsanopoulos, A. N., "Efficient Implementations of Quadratic Digital Filters", IEEE Trans. On ASSP, vol. 34, n^o 6, december 1986.
- [12] Chui, C. K., "An Introduction to Wavelets", Academic Press, New York, 1992.
- [13] Chui, C. K., "A Tutorial in Theory and Applications", Academic Press, New York, 1992.
- [14] Claussi, D., Jernigan, M., "Towards a Novel Approach for Segmentation of SAR Sea Ice Imagery", 26th symposium on Remote Sens. Of Environment and 18th Annual Symp. Of the Canadian Remote Sens. Society, 1995.
- [15] Connors, R. W., Harlow, C. A., "A Theoretical Comparison of Texture Algorithms", IEEE Trans. Pattern Anal. And Machine Intell., vol. 2, n^o 3, pp. 204-221, 1980
- [16] Croisier, A., Esteban, D. and Galand, C., "Perfect Channel Splitting by Use of Interpolation/Decimation/Tree Decomposition Techniques", Int. Conf. Inform. Sci. Syst., Patras, Grécia, 1976.
- [17] Daubechies, I., "Ten Lectures on Wavelets", Notes from the 1990 CBMS-NSF Conference on Wavelets and Applications at Lowell, SIAM, MA. Philadelphia, PA, 1992.
- [18] Daubechies, I., "Orthonormal Bases of Compactly Suported Wavelets", Comm. Pure Appl. Math., Vol. 41, pp 909-996, 1988.
- [19] Daubechies, I., "Wavelets and Other Phase Space Localization Methods", Proc. of ICM 94, Birkhäuser Verlag, Switzerland, 1995.

- [20] De Valois, R. L., Albrecht, D. G. and Thorell, L. G., "Spatial-frequency selectivity of cells in macaque visual cortex, *Vision Res.* Vol. 22, pp. 545-559, 1982
- [21] Donoho, D., "Abstract Statical Estimation and Modern Harmonic Analysis", *Proc. of ICM 94*, Birkhäuser Verlag, Switzerland, 1995.
- [22] Du Buf, J. M. H., Kardan, M., Spann, M., "Texture Feature Performance for Image Segmentation", *Pattern Recognition*, vol. 23, n^o 3/4, pp. 291-309, 1990.
- [23] Duda, R. O. and Hart, P. E., "Pattern Classification and Scene Analysis", *Jonh Wiley & Sons*, New York, 1973.
- [24] Dudgeon, D. E., Mersereau, R. M., "Multidimensional Digital Signal Processing", *Prentice Hall, Inc.*, New Jersey, 1984.
- [25] Dutra, L. V., Huber, R., "Feature Extraction and Selection for ERS-1/2 In SAR Classification", *Int. J. of Remote Sensing*, Vol. 20, n^o 5, 1999
- [26] Farrokhnia, F., "Multi-Channel Filtering Techniques for Texture Segmentation and Surface Quality Inspection. PhD thesis, Michigan State University, 1990.
- [27] Fu, S. K., "Syntatic Pattern Recognition and Appications", *Prentice Hall*, 1982.
- [28] Gabor, D., "Theory of Communication", *J. IEE*, vol. 93, pp. 429-457, 1946.
- [29] Gomes, J. e Velho, L., "Computação Gráfica: Imagem", *IMPA-SBM*, Rio de Janeiro, 1994.

- [30] Grossman, A., and Morlet, J., "Decomposition of Hardy functions into square integrable wavelets of constant shape", *SIAM J. Math.*, vol. 15, pp. 723-736, 1984
- [31] Haar, A., "Zur Theorie der Orthogonalen Funktionen-Systeme", *Math. Ann.* Vol. 69, pp 331-371, 1910.
- [32] Hager, William W., "Applied Numerical Linear Algebra", Prentice Hall College Div., 1987.
- [33] Haralick, R. M., Shanmugam, K. and Dinstein, I., "Textural features for image classification", *IEEE Trans. On SMC*, vol 3, pp. 610-621, 1973.
- [34] Haralick, R. M., "Statistical and Strutural Approches to Texture", *Proc. of IEEE*, vol 67(5), pp. 786-804, 1979.
- [35] Jain, A. K. and Bhattacharjee, S., "Text Segmentation Using Gabor Filters for Automatic Document Processing". *Machine Vision and Applications*, vol 5(3), pp. 169-184, 1992.
- [36] Ji, Q., Engel, J., Craine, E., "Texture Analysis for Classification of Cervix Lesions", *IEEE Trans. On Medical Imaging*, Vol. 19, n^o 11, Nov. 2000.
- [37] Julez, B., "Experiments in the Visual Perception of Texture", *Scientific American*, vol 212, 34-43, 1975
- [38] Julez, B., "Visual Pattern Discrimination", *IRE Transactions Inf. Theory*, vol 8, 84-92, 1962.
- [39] Julez, B.(b), "Textons, the Elements of Texture Perception, and their Interactions", *Nature(London)*, vol 290, 91-97, 1981.
- [40] Julez, B.(a), "A Theory of Preattentive Texture Discrimination Based on First-order Statistics of Textons", *Biol. Cybernectis*, vol 41, 131-138, 1981.

- [41] Laws, K. I., "Rapid Texture Identification", Proc. SPIE Conf. Image Processing for Missile Guidance, pp. 376-380, 1980.
- [42] Luenberger, D. G., "Introduction to Linear and Nonlinear Programming", Addison-Wesley,
- [43] Mallat, S. G., "Review of Multifrequency Channel Decomposition of Images and Wavelet Models", Technical Report n° 412, NYU, 56 pag., 1988.
- [44] Mallat, S., and Zhong, S. "Characterization of Signals from Multiscale Edges" *IEEE Trans. on PAMI*, vol. 14, no. 7, 1992
- [45] Mallat, S., and Zhong, S. "Wavelet Maxima Representation".in "Wavelets and Applications"- Proceedings of Int. Conf., Marseille, France, May 1989. Yves Meyer Ed., Masson, 1992.
- [46] Mallat, S.G., "A Wavelet Tour of Signal Processing", Academic Press, 1998.
- [47] Mallat, S., "A Theory for Multiresolution Signal Decomposition: the Wavelet Representation", *IEEE Trans. Pattern Anal. Machine Intell.*, vol 11(7), pp. 674-693, 1989.
- [48] Mallat,S., Zhang, Z., "Matching Pursuits with Time-frequency Dictionaries" *IEEE Trans. On Signal Processing*, Vol. 41, n° 12, pp 3397-3415, 1993.
- [49] Marple, S. L. Jr, "Digital Spectral Analysis With Applications", Prentice Hall, Inc., New Jersey, 1987.
- [50] Marr, D., and Hildreth, E., "Theory of Edge Detection", Proc. Royal Soc. London, vol. 207, pp. 187-217, 1980.
- [51] Mertzios, B. G., Sicuranza, G. L. and Venetsanopoulos, A. N., "Efficient Realization of Two-Dimensional Quadratic Digital Filters". *IEEE Trans. On ASSP*, vol. 37, n° 5, may 1989.

- [52] Meyer, Y. "Ondelettes et Opérateurs", Tome 1, Hermann, Paris, 1990.
- [53] Morlet, J., "Sampling Theory and Wave Propagation", in NATO ASI Series, Vol. 1, Issues in Acoustic/Image Processing and Recognition, C. H. Chen, ed., Springer-Verlag, pp 233-261, 1983.
- [54] Mudigonga, N. R., Rangayyan, R. M., Desautels, J. E. L., "Gradient and Texture Analysis for Classification of Mamographic Masses", IEEE Trans. On Medical Imaging, Vol. 19, n^o 10, Oct. 2000.
- [55] Novo, E. M. L. M., "Sensoriamento Remoto, Princípios e Aplicações", Ed. Edgard Blücher, 1989.
- [56] Nowak, R. D., "Volterra Filter Identification Using Penalized Least Squares", Proceedings of ICASSP 96, Atlanta, may, 1996.
- [57] Nowak, R. D. and Baraniuk, R. G., "Wavelet-Based Transformations For Nonlinear Signal Processing", IEEE Trans. On Signal Processing,
- [58] Ohanian, P. P., Dudes, R. C., "Performance Evaluation for Four Classes of Textural Features", Pattern Recognition, Vol. 25, n^o 8, pp 819-833, 1992.
- [59] Ojala, T., Pietikäinen, M., Harwood, D., "A Comparative Study of Texture Measures with Classification Based on Feature Distributions", Pattern Recognition, vol. 29, n^o 1, pp. 51-59, 1996.
- [60] O'Sullivan, F., "A Statistical Perspective on Ill-posed Inverse Problems", Statistical Science, vol. 1, n^o 4, pp 502-527, 1986.
- [61] Pagamisse, A., Dutra, L. V., "Segmentação de Texturas pela Transformada Wavelet", Série em Ciências Cartográficas, Vol. 2, pp , Editora da Univ. Federal de Curitiba, 2002.

- [62] Patel, D., "Page Segmentation for Document Image Analysis Using a Neural network. Optical Engineering, vol 35(7), pp. 1854-1861, 1996.
- [63] Pattichis, M. S., Pattichis, M. A., Bovik, A., Kyriacou, K., "AM-FM Texture Segmentation in Electron Microscopic Muscle Imaging", IEEE Trans. On Medical Imaging, Vol. 19, n^o 12, Dec. 2000.
- [64] Pietikäinen, M., Rosenfeld, A., Davis, L. S., "Experiments with Texture Classification Using Averages of Local Pattern Matches", IEEE Trans. Syst., Man, Cyb., vol. 13, n^o 3, pp. 421-426, 1983.
- [65] Pitas, I. and Venetsanopoulos, A. N., "Nonlinear Digital Filters –Principles and Applications", Kluwer Academic Publishers, 1990.
- [66] Ramponi, G. F., "Bi-Impulse Design of Isotropic Quadratic Filters", Proceedings of the IEEE, vol 78, n^o 4, april 1990.
- [67] Ramponi, G. F., Sicuranza, G. L. and Ukovich, W., "A computational Method For the Design Of 2-D Nonlinear Volterra Filters", IEEE Trans. On Circuits And Systems, vol. 35, n^o 9, 1988.
- [68] Ramponi, G. F., Sicuranza, G. L., "Quadratic Digital Filters For Image Processing", IEEE Trans. On ASSP, vol. 36, n^o 6, june 1988.
- [69] Ramponi, G. "Edge Extraction by a Class of Second-order Nonlinear Filters", Electronics Letters, vol 22(9), 1986
- [70] Rao, R. A. and Schunck, B. G., "Computing oriented texture fields. CVGIP: Graphical Models and Image Processing, vol 53(2), pp. 157-185, 1991.
- [71] Randen and J. H. Husøy, "Filtering for Supervised Texture Segmentation: A Comparative Study," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 21, n^o. 4, pp. 291-310, 1999. (<http://www.ux.his.no/~tranden/>)

- [72] Richards, J. A., Jia, X., "Remote Sensing Digital Image Analysis", Springer, Berlin, 1999.
- [73] Richards, W. and Polit, A., "Texture Matching", Kybernetic, vol. 16, pp. 155-162, 1974.
- [74] Sicuranza, G. L., "Quadratic Filters For Signal Processing", Proceedings of the IEEE, vol 80, n^o 8, august 1992.
- [75] Sklansky, J. "Image Segmentation and Feature Extraction", IEEE Trans. Man Cybern., vol. 8, pp. 237-247, 1978
- [76] Strand, J., Tax, T., "Local Frequency Features for Texture Classification", Pattern Recognition, Vol. 27, n^o 10, pp 1397-1406, 1994.
- [77] Super, B. J., and Bovik, A. C., "Shape from Texture Using Local Spectral Moments". IEEE Trans. Pattern Anal. And Machine Intell., vol 17(4), pp. 333-343, 1995.
- [78] Sweldens, W., "Wavelets: What Next ?", Proceedings of the IEEE, Vol. 84, n^o 4, pp 680-685, 1996.
- [79] Tamura, H., Mori, S. and Yamawaki, Y., "Textural Features Corresponding to Visual Perception", IEEE Trans. SMC, pp. 460-473, 1993.
- [80] Tuceryan, M. and Jain, A. K. "Texture Analysis" in Handbook of Pattern Recognition and Computer Vision, World Scientific Publishing Company, pp. 610-621, 1973.
- [81] Unser, M. and Eden, M., "Nonlinear Operators for Improving Texture Segmentation Based on Features Extracted by Spatial Filtering", IEEE Trans. On Systems, Man, and Cybernetics, vol. 20, n^o 4, 1990.

- [82] Unser, M., "Texture Classification and Segmentation Using Wavelet Frame", IEEE Trans. On Image Processing, vol. 4, n^o 11, Nov. 1995.
- [83] Vetterli, M., Kovacevic, J., "Wavelets and Subband Coding", Prentice Hall, New Jersey, 1995.
- [84] Walsh, J. L., "A Closed Set of Normal Orthogonal Functions", Am. J. Math., vol. 45, n^o 1, pp. 5-24, 1923.
- [85] Weldon, T. P., Higgins, W. E. and Dunn, D. F., "Efficient Gabor Filter Design for Texture Segmentation", Pattern Recognition, Vol. 29, N^o. 12, pp. 2005-2015, Dec. 1996.
- [86] Weszka, J. S., Dyer, C. R., Rosenfeld, A., "A Comparative Study of Texture Measures for Terrain Classification", IEEE Trans. Syst., Man, Cyb., Vol. 6, n^o 4, April 1976.
- [87] Witkin, A. P., "Recovering Surface Shape and Orientation from Texture", Artif. Intell., Vol. 17, pp 17-45, 1981.
- [88] Zucker, S. W., Kant, K., "Multiple-Level Representations for Texture Discrimination", Proc. IEEE Conf. On Pattern Recognition and Image Processing, Dallas, Texas, pp. 609-614, 1981.
- [89] MIT Vision Texture Database, "[<http://www-white.media.mit.edu/vismod/imagery/VisionTexture/vistex.html>]
- [90] http://www.cssip.uq.edu.au/meastex/www/for_images.html#meastex

Apêndice A

Filtros de Volterra

Filtros lineares são largamente utilizados em Processamento de Sinais devido a suas propriedades e simplicidade, além de apresentarem resultados satisfatórios em um grande número de aplicações. Entretanto, existem algumas situações em que os filtros lineares não dão bons resultados ou então não modelam determinados fenômenos adequadamente, por exemplo, na presença de ruídos dependentes do sinal ou multiplicativos e em sinais com estatísticas não-Gaussianas.

A fim de compensar as limitações dos filtros lineares, desenvolveram-se inúmeros filtros não-lineares. Sua motivação está no fato que na modelagem de sistemas, pode ser necessário levar em conta não-linearidades inerentes ou mesmo que caracterizam sistemas do mundo real. Em Processamento de Imagens é sabido que filtros lineares não são capazes de remover o ruído, em particular do tipo impulsivo superimposto, em uma imagem, sem borrar as bordas. Além disso, é freqüentemente necessário levar em conta o comportamento não-linear intrínseco do sistema visual humano ou de sistemas óticos, resultante da relação quadrática entre a intensidade ótica e o campo ótico.

As mais importantes famílias de filtros não-lineares são [65]:

- filtros homomórficos
- filtros morfológicos
- filtros de medianas
- filtros de ordens estatísticas (order statistics)

- filtros adaptativos
- filtros polinomiais ou de Volterra

Os filtros polinomiais são baseados em uma relação entrada-saída expressa na forma de uma série de Volterra discreta. Na prática, a série de Volterra pode ser considerada como uma série de Taylor com memória.

A fim de se obter uma representação notacionalmente e computacionalmente mais tratável, costuma-se truncar a série de Volterra, considerando-se apenas termos até segunda ordem. Os filtros resultantes são chamados de quadráticos. Assim, um filtro de Volterra quadrático bidimensional tendo suporte finito é definido como,

$$y(n_1, n_2) = h_0 + \bar{h}_1[X(n_1, n_2)] + \bar{h}_2[X(n_1, n_2)] \quad (7.1)$$

onde h_0 é um termo constante, \mathbf{X} é um sinal bidimensional,

$$\bar{h}_1[X(n_1, n_2)] = \sum_{k_1=0}^{N_1-1} \sum_{k_2=0}^{N_2-1} h_1(k_1, k_2) X(n_1 - k_1, n_2 - k_2) \quad (7.2)$$

é a parte linear do filtro não-linear, isto é, um filtro FIR bidimensional, e

$$\begin{aligned} \bar{h}_2[X(n_1, n_2)] = & \sum_{k_1=0}^{N_1-1} \sum_{k_2=0}^{N_2-1} \sum_{k_3=0}^{N_1-1} \sum_{k_4=0}^{N_2-1} h_2(k_1, k_2, k_3, k_4) X(n_1 - k_1, n_2 - k_2) \\ & X(n_1 - k_3, n_2 - k_4) \end{aligned} \quad (7.3)$$

é o operador quadrático de Volterra.

A fim de facilitar o manuseio de (7.1), buscou-se uma descrição matricial de (7.2) e (7.3). Isso é feito organizando-se os dados de entrada e os coeficientes lineares em duas matrizes X_1 e H_1 , enquanto que para o termo quadrático, o array X_2 pode ser obtido pelo produto de Kronecker $X_2 = X_1 \otimes X_1$, e os coeficientes quadráticos podem ser organizados em um array H_2 . Com essas substituições, (7.1) torna-se

$$y(n_1, n_2) = \text{tr}\{H_1 X_1^T\} + \text{tr}\{H_2 X_2^T\}$$

onde o expoente T indica transposição de matriz e $\text{tr}\{\}$ indica o traço de uma matriz.

No caso bidimensional, considera-se regiões quadradas de tamanho $N \times N$, isto é matrizes, e nesse caso, H_2 e X_2 são matrizes de tamanho $N^2 \times N^2$ ordenadas lexicograficamente em N^2 blocos quadrados de N^2 termos. Por exemplo, para $N=3$, tem-se:

$$H_2 = \begin{array}{|ccc|cc|} \hline h_{2L}(1,1) & h_{2L}(1,2) & h_{2L}(1,3) & & \\ h_{2L}(1,4) & h_{2L}(1,5) & h_{2L}(1,6) & h_{2L}(2,j) & h_{2L}(3,j) \\ h_{2L}(1,7) & h_{2L}(1,8) & h_{2L}(1,9) & & \\ \hline & h_{2L}(4,j) & & h_{2L}(5,j) & h_{2L}(6,j) \\ & h_{2L}(7,j) & & h_{2L}(8,j) & h_{2L}(9,j) \\ \hline \end{array}$$

Filtros de Volterra e, portanto, filtros quadráticos, possuem duas importantes propriedades [74]. A primeira, é que a saída do filtro de Volterra é linearmente dependente dos coeficientes do filtro. A segunda resulta da representação de não-linearidades do sinal por meio de operadores multidimensionais atuando em produtos das amostras do sinal discreto de entrada. A primeira característica é usada a fim de analisar filtros quadráticos, encontrar novas realizações, derivar adaptações de algoritmos, etc. A segunda permite a descrição do comportamento do filtro no domínio de frequência por meio de um tipo de convolução multidimensional.

Em suma, a expansão discreta de Volterra pode ser considerada como uma descrição matemática que, estendendo a clássica abordagem linear, é muito útil para a análise e síntese de sistemas não-lineares. De fato, ela define uma classe de filtros não-lineares, incluindo os filtros quadráticos como sua mais simples subclasse.

Várias aplicações de filtros quadráticos podem ser encontradas na literatura, envolvendo sinais discretos e sistemas, em uma ou mais dimensões. Para citar algumas delas [65], [74],

- modelagem de sistemas não-lineares
- cancelamento de eco (não-linear)
- cancelamento de ruído (adaptativo)
- detecção e estimação

- extração de bordas
- suavização com preservação de bordas [67]
- discriminação de texturas [66]
- predição e interpolação de seqüências de imagens de TV
- restauração de imagens [68]

Um dos aspectos mais importantes da filtragem quadrática está na procura de técnicas eficientes para a obtenção dos coeficientes dos filtros. Uma das maiores limitações está no fato que o problema de estimação requer a solução de sistemas de equações lineares que freqüentemente são mal condicionados [60]. Simplificações são possíveis e necessárias, dependendo do problema a ser modelado [67], [11]. Tais dificuldades têm levado autores a considerar abordagens do tipo,

- estruturas ortogonais para os filtros de Volterra
- lattice filter structures
- pseudo-random multilevel sequences
- decomposição paralela do filtro de Volterra
- mínimos quadrados penalizados
- wavelet-based transformation for nonlinear signal processing [57]

Uma das qualidades mais apregoadas sobre os filtros de Volterra é a sua superioridade em relação aos filtros lineares quando se procura extrair bordas de uma imagem na presença de ruídos. Em [69], Ramponi desenvolveu um conjunto de filtros de Volterra a fim de extrair bordas em presença de ruídos. Aplicamos este filtro a uma imagem corrompida com ruído impulsivo e comparamos o resultado dessa filtragem com o filtro clássico de Sobel para extração de bordas. A figura A.1 apresenta o resultado visual dessa filtragem.

obtidos a partir de amostras das texturas. Para obter esses coeficientes deve-se resolver sistemas não-lineares que conforme exposto acima são em geral mal-condicionados. A fim de superar esse mal-condicionamento faz-se necessário colocar algumas restrições no modelo, tais como, isotropia ou direcionalidade, de modo a obter convergência.

Apêndice B

Rotinas e programas utilizados na implementação do modelo experimental utilizando o Software IDL.

- Programa Principal:

```
pro geral,a,niveis,k,suav,bmax,saida
```

```
;Programa para a segmentação da imagem a, formada por k texturas.  
;As amostras colhidas serão de tamanho tam=32x32, bastando clicar na posição  
;(x,y) inicial para a escolha da amostra.  
;tam=32x32 pode ser alterado dependendo do problema.  
;suav e' o tamanho da região para cálculo da energia.  
;saida é o resultado da segmentação  
;bmax são as bandas de máxima separabilidade
```

```
t=size(a)  
tam=32  
n=k*(k-1)/2  
am=fltarr(tam,tam,k,4)  
x=intarr(k) & y=x  
temp=bytarr(n)  
Dgeral=fltarr(niveis,k,k,4)  
fisher=fltarr(niveis,n)  
b=fltarr(niveis,t(1),t(2),4)  
bmax=fltarr(niveis,t(1),t(2),n)  
Emax=fltarr(niveis,k,k,2)  
z=fltarr(t(1),t(2),n)
```

```
; Seleccionando as amostras  
tvsc1,a
```

```

print, '
for i=0,k-1 do begin
print,'posicao inicial da textura',i+1
cursor,p1,p2,/device
x(i)=p1 & y(i)=p2
print,x(i),y(i)
print, '

```

```

wait,1
endfor

```

; Transformada Wavelet da imagem a no nível v, separação das direções
; b2,b3,b4,b5, tomada do valor absoluto e homogeneização.

```

For v=1, niveis do begin
b2=fltarr(t(1),t(2)) & b3=b2 & b4=b2 & b5=b2
WT=bancof2(a,v)
WT=abs(WT)
b2(*,*)=WT(1,*,*) & b3(*,*)=WT(2,*,*)
b4(*,*)=WT(4,*,*) & b5(*,*)=WT(5,*,*)

```

```

k2=suav
b2=bordo(b2,k2) & b3=bordo(b3,k2)
b4=bordo(b4,k2) & b5=bordo(b5,k2)

```

```

aux=fgauss1(k2-1) & aux1=transpose(aux)
b2=convol(b2,aux) & b2=convol(b2,aux1)
b3=convol(b3,aux) & b3=convol(b3,aux1)
b4=convol(b4,aux) & b4=convol(b4,aux1)
b5=convol(b5,aux) & b5=convol(b5,aux1)

```

```

b2=b2(k2:t(1)+k2-1,k2:t(2)+k2-1) & b3=b3(k2:t(1)+k2-1,k2:t(2)+k2-1)
b4=b4(k2:t(1)+k2-1,k2:t(2)+k2-1) & b5=b5(k2:t(1)+k2-1,k2:t(2)+k2-1)

```

```

b(v-1,*,*,0)=b2(*,*) & b(v-1,*,*,1)=b3(*,*)
b(v-1,*,*,2)=b4(*,*) & b(v-1,*,*,3)=b5(*,*)

```

; Rotação Canônica

```

for p1=0,k-1 do for p2=0,3 do am(*,*,p1,p2)=b(v-1,x(p1):x(p1)+tam-1,
y(p1):y(p1)+tam-1,p2)

```

```

aux=0
for i=0,k-2 do for j=i+1,k-1 do begin
d=canonic4(am(*,*,i,0),am(*,*,i,1),am(*,*,i,2),am(*,*,i,3),am(*,*,j,0),am(*,*,j,1),am(*,*,j,2),am(*,*,j,3))
Dgeral(v-1,i,j,*)=transpose(d)

```

```

a1=d(0)*b2+d(1)*b3+d(2)*b4+d(3)*b5
a1=a1-min(a1)
a1=a1/max(a1)*255

```

```

;Cálculo das energias locais para cálculo dos coef. de Fisher.

```

```

kb=suav
a2=bordo(a1,kb) & a2=smooth(a2,kb) & a2=a2(kb:t(1)+kb-1,kb:t(2)+kb-1)
;a2=bordo(a1,kb) & a2=median(a2,kb) & a2=a2(kb:t(1)+kb-1,kb:t(2)+kb-1)
bmax(v-1,*,*,aux)=a2

```

```

am1=a2(x(i):x(i)+tam-1,y(i):y(i)+tam-1)
am2=a2(x(j):x(j)+tam-1,y(j):y(j)+tam-1)

```

```

m1=moment(am1) & m2=moment(am2)
fisher(v-1,aux)=(m1(0)-m2(0))^2/(m1(1)^2+m2(1)^2)
Emax(v-1,i,j,0)=m1(0)
Emax(v-1,i,j,1)=m2(0)

```

```

aux=aux+1
endfor
endfor

```

```

;Escolha dos coeficientes de Fisher.

```

```

aux=0
for i=0,k-2 do for j=i+1,k-1 do begin
niv=sort(fisher(*,aux))

```

```

a2(*,*)=bmax(niv(niveis-1),*,*,aux)

```

```

x1=abs(a2-Emax(niv(niveis-1),i,j,0))
x2=abs(a2-Emax(niv(niveis-1),i,j,1))

```

```

x3=x1 gt x2
x1=x3 eq 0
x2=x3 eq 1
saida=i*x1+j*x2
z(*,*,aux)=saida
aux=aux+1
endfor

```

```

;Limiarização

```

```

for i=0,t(1)-1 do for j=0,t(2)-1 do begin
temp(*)=z(i,j,*)
temp1=histogram(temp,min=0,max=k-1)
temp2=sort(temp1)
saida(i,j)=temp2(k-1)

```

```

endfor
return
end

```

- Rotinas:

Rotina Canonic4 (citada no corpo da tese)

```

function canonic4, a1,a2,a3,a4,a5,a6,a7,a8

; Análise Canônica para 4 bandas, 2 amostras
; Entradas: Matrizes a1,a2,a3,a4=mesma amostra das bandas 1,2,3 e 4
;respectivamente.
;Matrizes a5,a6,a7,a8=outra amostra das bandas 1,2,3, e 4 respectivamente.

aux1=transpose([[vet1(a1)],[vet1(a2)],[vet1(a3)],[vet1(a4)]])
aux2=transpose([[vet1(a5)],[vet1(a6)],[vet1(a7)],[vet1(a8)]])
m1=float(total(aux1,2)/n_elements(a1))
m2=float(total(aux2,2)/n_elements(a2))
m=(m1+m2)/2
sigma1=(correlate(aux1,/covariance)+correlate(aux2,/covariance))/2
sigma2=(m1-m)#transpose(m1-m)+(m2-m)#transpose(m2-m)

; Resolvendo agora a equação no determinante...
; Ver William W. Hager, "Applied Numerical Linear Alg. pag 305
; pois sigma1 e' Simétrica definida positiva.

; Fazendo a decomposição de Cholesky de sigma1...
aux=sigma1
choldc,aux,p,/double
for i=0,3 do aux(i,i)=p(i)
aux(1,0)=0 & aux(2,0)=0 & aux(3,0)=0 & aux(2,1)=0 & aux(3,1)=0 & aux(3,2)=0
Chol=invert(aux,/double)##sigma2##transpose(invert(aux,/double))
; Simetrizando...
temporario=Chol
for i=0,3 do begin
for j=i,3 do Chol(i,j)=temporario(j,i)
endfor
endfor

; O problema a ser resolvido se torna: Chol(y)=lambda(y)

lambda=eigenql(Chol,/double)

```

```

; Escalonando o sistema [sigma2-lambda*sigma1]d=0 e obtendo
; as relações entre as variáveis.
aux=sigma2-lambda(0,0)*sigma1
; Escalonando...
for i=0,1 do begin
  for j=i+1,3 do begin
    for k=i+1,3 do aux(k,j)=aux(i,i)*aux(k,j)-aux(i,j)*aux(k,i)
  endfor
endfor

; Usando a condição (10.18)
d3=-aux(3,2)/aux(2,2)
d2=-((aux(2,1)*d3+aux(3,1))/aux(1,1)
d1=-((aux(1,0)*d2+aux(2,0)*d3+aux(3,0))/aux(0,0)

temp1=sigma1(0,0)*d1^2+sigma1(1,1)*d2^2+sigma1(2,2)*d3^2+sigma1(3,3)
temp2=(sigma1(1,0)+sigma1(0,1))*d1*d2+(sigma1(0,2)+sigma1(2,0))*d1*d3+(sigma1(3,0)+sigma1(0,3))*d1
temp3=(sigma1(2,1)+sigma1(1,2))*d2*d3+(sigma1(3,1)+sigma1(1,3))*d2+(sigma1(3,2)+sigma1(2,3))*d3
temp=temp3+temp2+temp1
d4=sqrt(1./temp)
d3=d3*d4
d2=d2*d4
d1=d1*d4

return,d=[d1,d2,d3,d4]
end

```

Rotina Bordo

(necessária para resolver o problema da filtragem nas bordas laterais, inferior e superior. Supõem-se simetria).

```
function bordo, x,n
```

```

;OBS: o tamanho de n não pode ser maior que o tamanho de x
t=size(x)
col=t(1)
lin=t(2)
a1=x(1:n,*)
a2=x(col-(n+1):col-2,*)
a1=rotate(a1,-3)

```

```

a2=rotate(a2,-3)
z=[a1,x,a2]
a1=z(:,1:n)
a2=z(:,lin-(n+1):lin-2)
a1=rotate(a1,-1)
a2=rotate(a2,-1)
return,[[a1],[z],[a2]]
end

```

Rotina fgauss1

```

function fgauss1, m

; Filtro Gaussiano conforme livro, Jonas Gomes, Luis velho:
; "Computação Gráfica: Imagem", pags 204-205. – Caso 1D.

box=[1,1]
aux=float([0,box,0])
for i=2,m do begin
    aux=convol(aux,box,0,/edge_wrap)
    aux=[aux,0]
endfor

return, double(aux(1:m+1)/2.^m)
end

```

Rotina bancof2

```

; Decomposição 2-dimensional por wavelet diádicas direcionais.
; Entradas: matriz nxn, e nível de decomposição.
; Saída: matriz b[6,m,n] onde
; b1= banda LL, b2=banda H(coluna), b3=banda H(linha), b4=banda HDHD,
; b5=banda DH(principal), b6=banda DH(secundaria) e D e' a rotação de 45 graus

```

```

function bancof2, imagem,niv

t=size(imagem)
g=[1,-1]

h=[.125,.375,.375,.125]
k=4
s=float(imagem)

```

```

sd=s

for nivel=1,niv do begin

k=2^(nivel-1)-1
t1=size(h)
h1=fltarr(t1(1)+k*(t1(1)-1))
for i=0,t1(1)-1 do h1(i+k*i)=h(i)
k1=size(h1) & k1=k1(1)
t1=size(g) & g1=fltarr(t1(1)+k*(t1(1)-1))
for i=0,t1(1)-1 do g1(i+k*i)=g(i)
k2=size(g1) & k2=k2(1)

hd=fltarr(k1,k1)
for i=0,k1-1 do hd(i,i)=h1(i)

gd=fltarr(k2,k2)
for i=0,k2-1 do gd(i,i)=g1(i)

; Filragem
k=max([k1,k2])
s=bordo(s,k)
sd=bordo(sd,k)

aux11=convol(s,h1) & h1=transpose(h1) & aux11=convol(aux11,h1)
aux12=convol(s,g1) & g1=transpose(g1) & aux13=convol(s,g1)

aux14=convol(sd,hd) & hd=rotate(hd,-1) & aux14=convol(aux14,hd)
aux15=convol(sd,gd) & gd=rotate(gd,-1) & aux16=convol(sd,gd)

aux11=aux11(k:t(1)+k-1,k:t(2)+k-1)
aux12=aux12(k:t(1)+k-1,k:t(2)+k-1)
aux13=aux13(k:t(1)+k-1,k:t(2)+k-1)
aux14=aux14(k:t(1)+k-1,k:t(2)+k-1)
aux15=aux15(k:t(1)+k-1,k:t(2)+k-1)
aux16=aux16(k:t(1)+K-1,k:t(2)+k-1)

s=aux11
sd=aux14
endfor

b=fltarr(6,t(1),t(2))
b(0,*,*)=aux11 & b(1,*,*)=aux12 & b(2,*,*)=aux13
b(3,*,*)=aux14 & b(4,*,*)=aux15 & b(5,*,*)=aux16

return, b
end

```