SEGMENTAÇÃO DE TEXTURAS PELA TRANSFORMADA WAVELET DIÁDICA DIRECIONAL

AYLTON PAGAMISSE[(1)]
LUCIANO VIEIRA DUTRA[(2)]

[1] Universidade Estadual Paulista - Unesp - Departamento de Matemática, Estatística e Computação, Campus de Presidente Prudente - aylton@prudente.unesp.br
[2] Instituto Nacional de Pesquisas Espaciais - INPE - Departamento de Processamento de Imagens - dutra@dpi.inpe.br

RESUMO

A segmentação de texturas tem um papel muito importante em sensoriamento remoto e em muitas outras áreas de aplicações. Seu estudo já se estende por mais de três décadas; contudo, continua sendo tópico de grande interesse. As pesquisas são motivadas em parte pela grande quantidade de texturas presentes em imagens e pela capacidade do sistema visual humano em lidar com elas. Apesar da grande quantidade de trabalhos sobre o assunto, segmentação de texturas continua apresentando ainda hoje, desafios e inúmeros problemas a serem resolvidos. Intuitivamente, textura está relacionada a variações tonais na imagem, que podem ser vistas como descontinuidades locais. Em imagens, descontinuidades estão associadas a bordas que geralmente são usadas para descrever a informação estrutural dos contornos de objetos. Pagamisse e Dutra (2002) desenvolveram um modelo para a segmentação de texturas a partir da representação de um sinal utilizando bordas em multiescala extraídas pela transformada *wavelet* diádica. Nesse trabalho, desenvolve-se uma forma mais geral da transformada *wavelet* diádica ao se incluir no modelo matemático e no modelo experimental as direções. Com esse aperfeiçoamento, melhora-se os resultados obtidos pelo modelo anterior, conforme pode ser verificado nas tabelas comparativas.

Palavras-chave: Textura, *Wavelet* diádica direcional, Segmentação.

TEXTURE SEGMENTATION BY ORIENTED DIADIC WAVELET TRANSFORM

ABSTRACT

Texture segmentation has a very important role in remote sensing and many other application areas. It has been studied for over three decades; however, it continues to be a very interesting topic. Research projects are motivated in part by the large number of textures present in images and by the ability of the human visual system in dealing with them. Despite extensive work published on this subject, texture segmentation still presents challenges and a number of problems to be solved. Intuitively, texture is related to tonal variations in the image, which can be seen as local discontinuities. In images, discontinuities are associated to edges that generally are used to describe structural information on the contours of the objects. Pagamisse & Dutra (2002) developed a model for the segmentation of textures starting from the representation of a signal using edges in multiscale extracted by the dyadic wavelet transform. In this work, we develop a more general way of the dyadic wavelet transform including in the mathematical and experimental models the directions. With this enhancement, the results are better than those obtained in the previous model, as shown in the comparative tables.

Keywords: *Texture, Oriented Diadic Wavelet, Segmentation.*

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 capacidade do sistema visual humano em lidar

com elas. Até o momento, o instrumento mais eficiente para detectar e discriminar 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 Julez (1962 e 1975), 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 terem elas 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 Julez 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.

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, quase-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 uma distância maior, essa textura pode ser totalmente diferente.

Em Tucerian e Jain (1973) encontram-se inúmeros exemplos de tentativas de definições de textura, mostrando que as diferentes definições dependem das aplicações, em particular, nas quais o conceito de textura é utilizado.

Intuitivamente, textura está relacionada a variações tonais (níveis de cinza) na imagem, sendo 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.

Devido à complexidade das estruturas que podem estar presentes em uma imagem, tais como bordas e texturas, faz-se necessário buscar uma representação da imagem que forneça uma caracterização de atributos que em princípio não estão claramente apresentados na representação original. Obviamente, essa representação dependerá do contexto e dos aspectos a serem realçados, ou seja, da finalidade e da fidelidade da representação. De modo geral, uma representação adequada é um pré-requisito para se obter técnicas de processamento eficientes, tais como realçamento, filtragem e restauração, análise, reconstrução etc, além disso, 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 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 (Mallat, 1998).

Nesse trabalho, a discriminação de texturas é obtida a partir da quantificação do conteúdo de bordas, 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, a fim de melhor discriminar as texturas 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 freqüência espacialmente orientados. Esta teoria foi proposta por Campbel e Robson (1968). Experimentos psicofisiológicos posteriores forneceram evidências adicionais como suporte a essa teoria (De Valois *et al*, 1982), sugerindo que existem mecanismos no córtex visual dos mamíferos que são sintonizados (*tuned*) a combinações de freqüência 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, desenvolve-se nesse trabalho 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 freqüência; além disso, em termos matemáticos, essa é uma representação completa e estável de um sinal discreto bidimensional.

Em Pagamisse e Dutra (2002), foi desenvolvido um modelo experimental utilizando a transformada *wavelet* diádica construída por Mallat e Zhong (1992). Nesse trabalho, desenvolve-se uma forma mais geral da transformada *wavelet* diádica ao se incluir no modelo matemático e no modelo experimental as direções aperfeiçoando o modelo anterior. Essa nova representação também é invariante por translação e se caracteriza por decompor um sinal num esquema multiresolucional.

2. TRANSFORMADA WAVELET DIÁDICA

Seja $\psi_s(x) = \frac{1}{s}\psi\left(\frac{x}{s}\right)$. A transformada *wavelet* (WT) de uma função *f(x)* na escala *s* e posição *x* é dada pela convolução:

$$W_s f(x) = f * \psi_s(x) = \int_{-\infty}^{+\infty} f(t)\frac{1}{s}\psi\left(\frac{x-t}{s}\right)dt \qquad (1)$$

A *wavelet* $\psi(x)$ pode ser interpretada como a resposta a impulso de um filtro passa banda e a WT como uma convolução com um filtro passa banda que é dilatado. Se a escala *s* é grande, $W_s f(x)$ detecta as componentes de baixa freqüência do sinal *f(x)*, enquanto que se a escala *s* diminui, o suporte de $W_s(x)$ diminui, de modo que a transformada *wavelet* $W_s f(x)$ se torna sensível a detalhes finos.

Grossman e Morlet (1984) 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.

A WT é chamada diádica quando o parâmetro de escala é discretizado ao longo da seqüência diádica $\left[2^j\right]_{j\in Z}$, ou seja, a WT de *f(x)* na escala 2^j e na posição *x* é dada pela convolução:

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

Se a *wavelet* ψ satisfizer certas condições (Mallat e Zhong, 1992), a função *f(x)* pode ser recuperada a partir de sua WT diádica.

A convolução apresentada em (2) é a versão contínua da WT diádica. Para fins computacionais, necessita-se de uma formulação discreta, além disso, é de interesse que se possa 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. Em Pagamisse e Dutra (2002), é apresentado uma família de *wavelets* diádicas, juntamente com sua discretização, construídas a partir de B-splines de ordem n, obtidas pela convolução de n splines de ordem zero. Nesse trabalho, utiliza-se essa mesma família de *wavelets*.

2.1 – IMPLEMENTAÇÃO NO CASO DISCRETO

A WT diádica discreta se caracteriza por decompor um sinal num esquema multiresolucional, invariante por translação e não realizar decimação nas saídas, diferentemente das implementações clássicas das *wavelets* de Daubechies e Mallat. Essas características são de fundamental importância para o sucesso do modelo de segmentação desenvolvido.

A WT diádica de um sinal em $\lambda^2(Z)$, isto é, sinais discretos de energia finita, é implementada como o operador linear limitado $\mathbf{W}: \lambda^2(Z) \to \lambda^2(I)$, onde, $I = \{1,2,...,J+1\} \times Z$, consistindo de J+1 operadores lineares

$$W_j: \lambda^2(Z) \to \lambda^2(Z),\ j = 1,2,...,J+1.$$

Os operadores W_j são operadores de convolução com filtros cujas respostas à impulso são dados por $(h_k)_{k\in Z}$ e $(g_k)_{k\in Z}$, e correspondem aos coeficientes das funções de transferência $\hat{H}(\omega)$ e $\hat{G}(\omega)$ (Pagamisse e Dutra, 2002). Na tabela 1, encontram-se valores de $(h_k)_{k\in Z}$ para alguns valores de n, e $(g_k)_{k\in Z}$ é da forma, $g_0 = -g_1 = 0.5$ e $g_k = 0$ para os demais $k \in Z$.

Se H_p e G_p são os filtros discretos obtidos ao colocar-se 2^p-1 zeros entre cada coeficiente dos filtros H= { h_k } e G={ g_k }, então, definindo $S_0f = f$, a decomposição *wavelet* diádica de um sinal discreto *f* é implementada da seguinte forma:

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

onde

$W_j f = W_{2^j} f$, $S_j f = S_{2^j} f$, $0 \le j \le J$ e (*) denota convolução.

A figura 1 apresenta de modo esquemático a decomposição de um sinal unidimensional pela WT diádica discreta. Nas componentes W_j encontram-se as altas freqüências do sinal, enquanto que as componentes $S_j f$ são versões suavizadas do sinal original nas várias escalas.

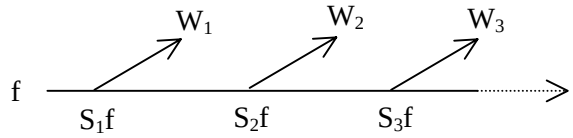

Figura 1 - Esquema de Decomposição 1D Pela WT Diádica.

Se o sinal original {$S_0f(n)$; $n \in Z$} tem N amostras, é possível computar no máximo $\log_2(N)$ escalas da transformada *wavelet* diádica discreta.

Tabela 1 - COEFICIENTES DA FUNÇÃO DE TRANSFERÊNCIA $\hat{H}(\omega)$

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

2.3 – TRANSFORMADA WAVELET DIÁDICA DE IMAGENS

A WT diádica pode ser estendida a $L^2(R^2)$ introduzindo-se várias *wavelets*, $\psi^k(x,y) \in L^2(R^2)$, $1 \le k \le K$, onde cada *wavelet* $\psi^k(x,y)$ tem uma orientação específica.

Seja

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

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

$$W f = (W^1_{2^j} f(x,y), ..., W^k_{2^j} f(x,y)).$$

Para assegurar que exista uma representação completa e estável de f(x,y), o plano de Fourier deve ser coberto pelas dilatações de $\hat{\psi}^k(\omega_x, \omega_y)$ (Mallat,1998). Isto significa que deve existir $A, B > 0$ tais que

$$\forall (\omega_x, \omega_y) \in R^2 - \{(0,0)\}, \quad A \le \sum_{k=1}^{K} \sum_{j=-\infty}^{+\infty} \left|\hat{\psi}^k(2^j\omega_x, 2^j\omega_y)\right|^2 \le B \ .$$

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

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

A figura 2 representa de modo esquemático uma família de seis *wavelets* com orientações diferenciadas.

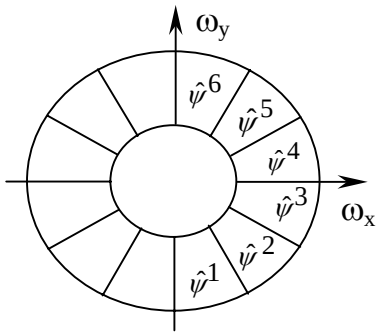

Figura 2 - Família de *Wavelets* com 6 Orientações

No modelo bidimensional, construído abaixo, serão utilizadas quatro orientações, θ = 0, 45, 90 e 135 graus. Para isso, define-se as quatro *wavelets* ψ^{1}, ψ^{2}, ψ^{3} e ψ^{4} como

$$\begin{aligned}
\psi^{1}(x,y) &= \psi(x)2\phi(y), \\
\psi^{2}(x,y) &= 2\phi(x)\psi(y), \\
\psi^{3}(x,y) &= \psi^{1}\left(\frac{\sqrt{2}}{2} R_{45}(x,y)\right), \\
\psi^{4}(x,y) &= \psi^{2}\left(\frac{\sqrt{2}}{2} R_{45}(x,y)\right)
\end{aligned} \qquad (3)$$

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

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

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

$$G(\omega_x - \omega_y) = \sum g_k \, e^{-ik(\omega_x - \omega_y)} = \sum g_k \, e^{-ik\omega_x} e^{ik\omega_y} \qquad (5)$$

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.

2.4 – IMPLEMENTAÇÃO DISCRETA BIDIMENSIONAL

As relações de escala em (4), juntamente com (5), permitem uma implementação rápida da transformada *wavelet* diádica. Para isto, seja D_1F a matriz cuja diagonal principal é formada pelos coeficientes do filtro F; D_2F 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:

$$W^1_{2^{j+1}} f = S^1_{2^j} f * (G_j . I)$$
$$W^2_{2^{j+1}} f = S^1_{2^j} f * (I . G_j)$$
$$W^3_{2^{j+1}} f = S^1_{2^j} f * (D_1 G_j . I)$$
$$W^4_{2^{j+1}} f = S^1_{2^j} f * (I . D_2 G_j)$$
$$S^1_{2^{j+1}} f = S^1_{2^j} f * (H_j . H_j)$$
$$S^2_{2^{j+1}} f = S^2_{2^j} f * (D_1 H_j . D_2 H_j)$$

onde I é o filtro de Dirac, cuja resposta a impulso é igual a 1 na posição zero, e zero nas demais posições; e (*) significa convolução.

Uma versão esquemática da implementação discreta do modelo de filtragem bidimensional com quatro direções é apresentada na figura 3 abaixo.

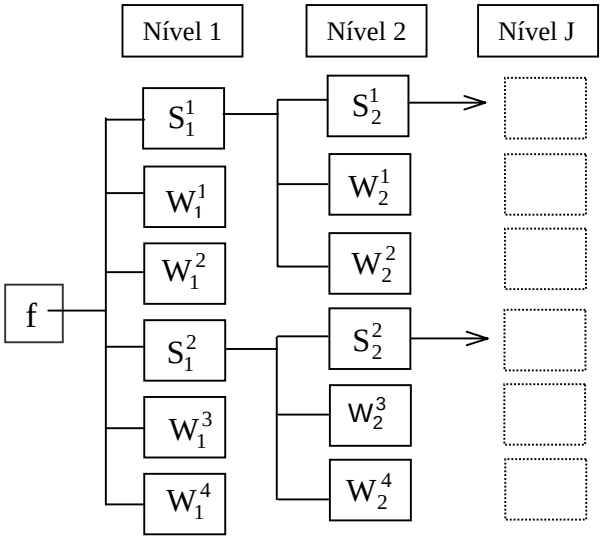

Figura 3: Esquema Da Decomposição Bidimensional Pela WT Diádica Com Quatro Direções.

3. MODELO EXPERIMENTAL

A hipótese desse trabalho é que medidas da distribuição da energia local, resultante da decomposição de uma imagem dada pela transformada *wavelet* diádica discreta nas suas várias escalas e orientações, são suficientes para discriminar inúmeras texturas. Isto decorre do fato que

$\left| W^k f(u,v,2^j) \right|^2$ é máximo quando o ângulo α_k casa a orientação dos componentes da textura e quando 2^j corresponde à escala dessas estruturas.

Com relação ao uso da energia local, pode-se destacar que esse procedimento é bastante utilizado não só no contexto da transformada *wavelet*, mas também no da transformada de Gabor ou mesmo no de matrizes de co-ocorrência, (Conners e Harlow,1980) (Dunn *et al*,1994) (Dunn e Higgins, 1995) (Haralick *et al*, 1973).

O modelo experimental apresentado nesse trabalho, busca agregar as respostas da transformada *wavelet*, nas escalas e orientações, de modo a encontrar as bordas de regiões de texturas homogêneas.

A maior dificuldade é que dentro de uma única região, para cada escala e orientação, a energia da *wavelet*, $\left| W^k f(u,v,2^j) \right|^2$, pode ter um alto grau de variabilidade. Desse modo, para fins de segmentação das texturas, faz-se necessário utilizar alguns procedimentos, tais como homogeneização, seleção e combinação das saídas mais adequadas.

Os procedimentos envolvidos no cálculo da energia local, no nível J da WT diádica direcional, são os seguintes:

- decomposição *wavelet* diádica
- tomada do valor absoluto das saídas da decomposição
- homogeneização
- combinação de atributos
- cálculo das energias locais
- seleção da banda de melhor separabilidade

Após a decomposição *wavelet*, as informações correspondentes a diferenças entre regiões de texturas estão concentradas em diferenças na variância, visto que todas as bandas possuem inicialmente média zero. O próximo passo é retificar as respostas dos filtros por aplicar-se uma transformação não linear. Esta operação tem o efeito de converter diferenças de variância em diferenças de média. A homogeneização subseqüente nos canais se insere no contexto de que textura é propriedade local e não pontual, ou seja, é indefinida para um único *pixel*, estando associada a um conjunto de *pixels*. Ela possibilita regiões de texturas homogêneas, além disso, reduz a variância, resultando em classes no espaço atributo que são mais compactas e conseqüentemente mais fáceis de distinguir.

A homogeneização pode ser realizada com vários tipos de filtros, como por exemplo o filtro Gaussiano passa-baixa,

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

Após a homogeneização, obtém-se quatro bandas por nível da transformada *wavelet*. Cada banda traz informações diferenciadas sobre a imagem, visto que são oriundas de orientações específicas. A fim de se obter a maior separabilidade possível entre as texturas, é feito uma combinação de atributos, que consiste em compor essas bandas e encontrar a direção canônica que dá a máxima discriminação (Duda e Hart, 1973). Feita a projeção nessas direções canônicas, tem-se em mãos quinze bandas. Deve-se então selecionar qual a melhor delas. Para isso, é utilizado o Critério de Fisher (Duda e Hart, 1973),

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

onde μ_1, μ_2, σ_1 e σ_2 são as médias e variâncias das amostras de duas texturas. 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 obtiver o maior índice no Critério de Fisher será gerada a banda de energia local e considerada como a banda de maior separabilidade no nível. A energia local é então tomada em regiões de mesmo tamanho que a amostra. Deve-se ter os mesmos cuidados na escolha do cálculo da energia que se teve ao escolher as amostras representativas. As regiões consideradas ao redor do *pixel* devem conter informação suficiente para caracterizar as texturas. Como esse procedimento é feito em amostras das texturas, tem baixo custo computacional.

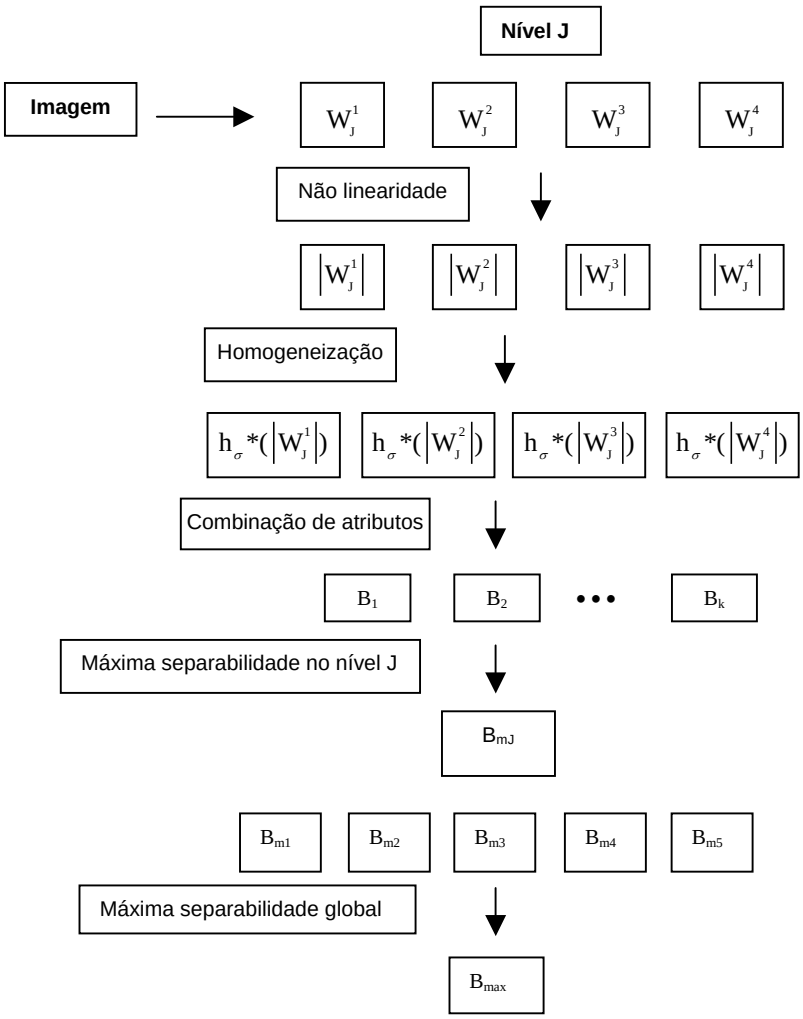

Figura 4 – Esquema Representativo do Modelo Experimental

3.1 – REALIZAÇÃO EXPERIMENTAL

Os passos necessários para a implementação experimental estão delineados no esquema apresentado na figura 4 e discriminados abaixo.

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, obtém-se quatro bandas, b_1, b_2, b_3 e b_4, conforme os filtros gerados pela implementação discreta da WT.
2- A partir dessas quatro bandas, obtém-se a banda com a máxima separabilidade entre as duas texturas, conforme descrito no item “combinação de atributos”.
3- Compara-se a discriminalidade em cada nível, a fim de verificar qual o melhor, obtendo-se os parâmetros finais que serão utilizados na imagem original.
4- Faz-se um threshold para segmentar as duas texturas.

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.

No caso de mais que duas texturas, aplica-se o mesmo procedimento para cada par de texturas e usa-se alguma técnica de segmentação. Nesse trabalho, utiliza-se o algoritmo combinatorial desenvolvido por Pagamisse e Dutra (2002).

4. APLICAÇÕES DO MODELO EXPERIMENTAL

A fim de avaliar a qualidade do modelo experimental desenvolvido, aplica-se em mosaicos de pares de texturas e de cinco texturas, comparando-se os resultados obtidos com métodos encontrados na literatura. Os resultados obtidos estão dispostos em tabelas comparativas.

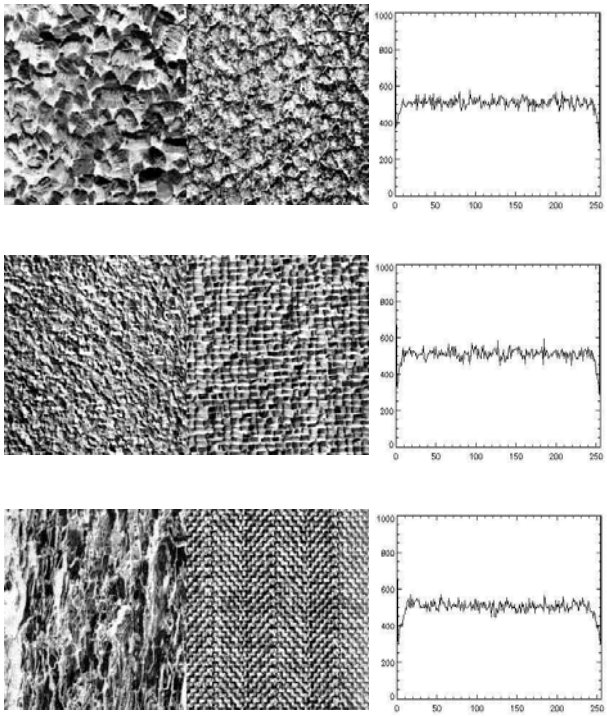

Figura 5 – Mosaicos D004D084, D012D017, D005D092 e Respectivos Histogramas.

A maior parte das texturas utilizadas na comparação de modelos de discriminação de texturas são extraídas do álbum “A Photografic Album for Artists and Designers” de Phil Brodatz (Brodatz,1966), que consiste 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 as utilizam como uma referência, pois representam um conjunto de texturas de fácil acesso e boa diversidade.

Não fugindo à regra, nesse trabalho optou-se por utilizar algumas dessas texturas a fim de avaliar a performance do método apresentado e compará-lo a outros métodos.

A comparação dos resultados é feita tomando-se como base o artigo de Randen e Husoy (1999). Provavelmente, este é o estudo mais completo publicado até hoje comparando abordagens baseadas em filtragem. Antes desse trabalho, haviam poucas comparações entre esquemas de extração de atributos de textura. Podemos citar: Weszka *et al* (1976), Conners e Harlow (1980), Du Buf *et al* (1990), Ojala *et al* (1996) e Strand e Tax (1994) entre outros. De modo geral, nesses trabalhos, atributos oriundos de matrizes de coocorrência também foram utilizados e mostraram-se superiores aos demais em alguns trabalhos mas não foram os melhores em outros, apresentando conclusões contrárias. Essa discrepância deve-se ao fato de não haver uma padronização nos procedimentos. Em Randen e Husoy (1999), encontramos uma extensa avaliação de abordagens relacionadas a atributos oriundos de filtragem. Nesse trabalho, os autores procuraram garantir uma certa padronização nos procedimentos, resguardadas suas particularidades, e os aplicaram a um conjunto de imagens do tipo Brodatz e outro do tipo MIT Vision Texture Database (http://vismod.media.mit.edu/vismod/imagery/VisionTexture/vistex.html) a fim de obter comparabilidade dos métodos. Além disso, devido ao seu alto grau de eficiência, popularidade e facilidade de implementação, atributos oriundos das matrizes de coocorrência também são utilizados nessa avaliação. Mais recentemente, Ojala *et al* (2001) também faz uso dos mesmos mosaicos de texturas para fins de comparação de perfomance de seu método. Na tabela 2 são apresentados os melhores resultados de Randen e Husoy, juntamente com os resultados obtidos pelas *wavelets* diádicas simples e direcionais.

4.1 – SEGMENTAÇÃO DE PARES DE TEXTURAS

Se obtivermos bons resultados na segmentação de pares de texturas, então estaremos aptos a reproduzir esses bons resultados quando nos depararmos com uma imagem que apresente uma quantia maior de texturas. Os pares de texturas da figura 5 são utilizados em Ojala *et al* e Randen e Husoy em

sua extensiva avaliação e os resultados obtidos pelos inúmeros métodos por eles avaliados, encontram-se apresentados na tabela 2.

A fim de que as imagens da figura 5 não sejam discrimináveis com simples estatísticas de primeira ordem, elas foram “homogeneizadas” com relação aos quatro primeiros momentos, isto é, todas possuem a mesma média, a mesma variância, a mesma kurtose e o mesmo *skewness*. Como resultado dessa homogeneização, pode-se observar como seus histogramas são praticamente idênticos e unimodais.

Na figura 6, mostramos a banda de máxima separabilidade do par de texturas D004D084 juntamente com sua superfície dos tons de cinza, seu histograma e o resultado da segmentação. Note que o histograma é bimodal, ressaltando a possibilidade de segmentação dos pares de texturas com um simples *trheshold*. No nosso caso,o *trheshold* foi feito a partir da energia média obtida em cada amostra, ou seja, de modo automático.

Os passos utilizados para se gerar das bandas de máxima separabilidade são aqueles descritos na seção 3 quando da descrição do modelo experimental. As amostras utilizadas são de tamanho 32x32 *pixels*, extraídas diretamente das imagens com o uso do mouse. As amostras devem refletir da melhor maneira possível o comportamento “textural” da imagem, isso é, devem ser tomadas de modo a estarem representando o comportamento de cada textura.

Nos pares apresentados na figura 5, podemos constatar 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.

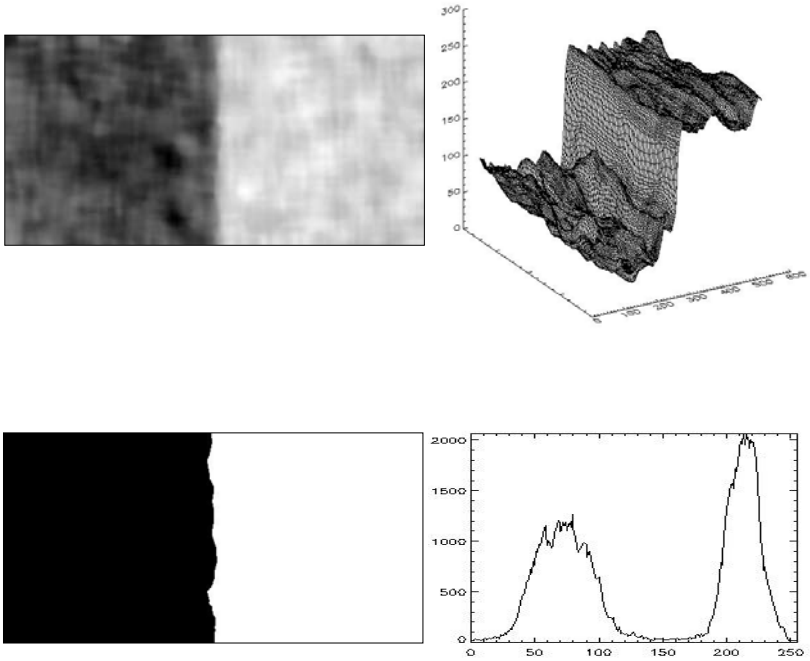

Figura 6 – Banda de Máxima Separabilidade para o Par de Texturas D004D084, Superfície de Tons de Cinza, Histograma e o Resultado da Segmentação.

4.1.1 - Discussão dos Resultados

O método aqui apresentado obteve o melhor resultado na segmentação do mosaico D005D092, praticamente empatou no mosaico D004D084 e foi inferior no mosaico D012D017. Note que a comparação é contra o melhor resultado entre quase 20 métodos. A diferença, em todos os casos, é pequena em relação aos métodos que apresentam os melhores resultados. Além disso, é interessante notar que o método utilizando as *wavelets* diádicas direcionais mostra-se superior ao das *wavelets* diádicas tradicionais.

De modo geral, 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 na superfície que aparecem na figura 6**.** Apenas para efeito de comparação, note que

uma única coluna num mosaico de duas texturas, representa 256 *pixels*, isto é, 0,39% do total de *pixels*. Assim, duas colunas e meia, mal classificadas, na borda entre as duas texturas, representam aproximadamente 1% de erro.

Podemos concluir, pelo menos nesses três pares de texturas que o método aqui apresentado obteve uma boa performance e, em média, foi melhor do que os outros. 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 bastante robusto.

4.2 - SEGMENTAÇÃO DE MOSAICOS DE CINCO TEXTURAS

Seguindo a apresentação de Randen e Husoy com relação ao tipo de testes efetuados, aplicamos o modelo experimental aqui desenvolvido para segmentar mosaicos com cinco texturas. Em problemas reais, dificilmente nos deparamos com uma grande quantidade de texturas que necessitem ser diferenciadas. Vamos utilizar esses mosaicos formados por cinco texturas, os quais foram obtidos da mesma forma que os pares de texturas utilizados inicialmente, isto é, por download do site http://www.ux.his.no/~tranden. A figura 7 apresenta esses mosaicos, enquanto que a tabela 2 mostra a performance de vários métodos utilizados para segmentá-los, conforme Randen e Husoy, juntamente com os resultados obtidos com as WT diádicas e diádicas direcionais. As texturas desses mosaicos não são oriundas exclusivamente de Brodatz, apenas a Nat5c, as demais foram obtidas de MIT Vision Texture Database (http://whitechapel.media.mit.edu/vismod/imagery/VisionTexture/vistex.html) e de (http://www.ee.oulu.fi/research/imag/texture/).

As amostras usadas a fim de 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 resultados obtidos por nosso modelo experimental, juntamente com resultados dos outros métodos, são mostrados na tabela 2.

Tabela 2 – Porcentagem de Erro para Métodos de Segmentação Aplicados aos Mosaicos das Figuras 5 e 7

Mosaico	**WT Diádicas 2 direções**	**WT Diádicas Direcionais**	**Melhor de Randen e Husoy**	
Nat5c	11,7	6,9	Opt. Repr. Gabor filter bank	7,2
Nat5v	17,7	9,5	F16b (d) (full rate)	18,9
Nat5v2	22,3	15,7	F_2_1_smpl (d) (Full rate)	20,6
Nat5v3	27,8	17,9	F32d (d) (full rate)	16,8
Nat5m	18,2	10,2	F16b (d) (full rate)	17,2
Média	**19,5**	**12,0**		**16,14**
---	---	---	---	---
D004D084	0,8	0,6	J_{MS}, J_U, J_F	0,7
D012D017	2,6	1,1	F32d (d) (full rate)	0,2
D005D092	2,4	1,1	DCT	2,5
Média	**1,9**	**0,9**		**1,1**

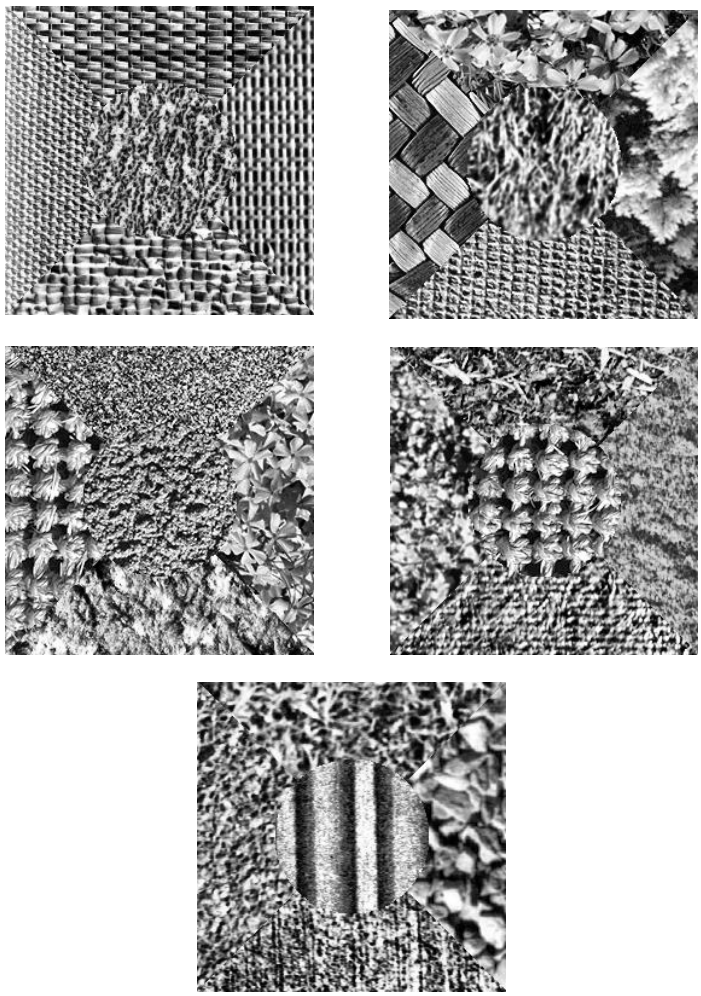

Figura 7 – Mosaicos Nat5c, Nat5v, Nat5v2, Nat5v3, Nat5m

4.2.1 - Discussão dos Resultados

Podemos observar a boa performance de nosso método quando comparado aos outros. Ele é inferior em apenas um caso e praticamente empata em outro. Nos demais, apresenta um resultado melhor. Interessante também realçar sua superioridade em relação ao método das *wavelets* diádicas tradicionais.mosaicos. Cabe dizer que os resultados que apresentamos são os melhores entre algumas tentativas. Tal procedimento é o mesmo que é utilizado por Handen e Husoy, isto é apresentam os melhores resultados obtidos. A figura 8 apresenta os resultados da segmentação dos mosaicos com cinco texturas na forma de uma imagem erro correspondente a cada mosaico segmentado.

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, enquanto que nos demais mosaicos encontramos uma variabilidade bem maior. Isso se deve à natureza homogênea das texturas em 5c e a não homogeneidade observada nos outros mosaicos. Nesses mosaicos, encontramos regiões que, pode-se dizer, não pertencerem à textura na qual estão contidas. É como se fossem “buracos” ou então, seus *texels* dentro dessas texturas são grandes. Esse comportamento influi não só na variabilidade mas também no percentual de erro. Essa alta variabilidade pode ser também observada não apenas em nosso método mas nos demais métodos apresentados, fazendo que a performance decaia de modo diretamente relacionado.

4.3 – CONSIDERAÇÕES FINAIS

Ao se considerar mais direções, além das tradicionais horizontal e vertical, na transformada *wavelet* diádica, aumenta-se a quantidade de informação a ser avaliada, fazendo que aparentemente essa nova formulação se torne computacionalmente mais pesada. Contudo, como todo procedimento é feito em pequenas amostras extraídas das texturas analisadas, o custo computacional continua baixo. Esse

aumento de informação, mostrou ser importante na medida em que melhorou os resultados de modo até significativo em alguns casos, mostrando a validade do esforço teórico e prático resultante.

Uma seqüência natural para esse trabalho será o de buscar formas de implementar essas idéias num esquema de segmentação não supervisionada, além disso, há a necessidade de utilizá-lo em situações mais realísticas a fim de conhecer suas limitações e buscar formas de superá-las.

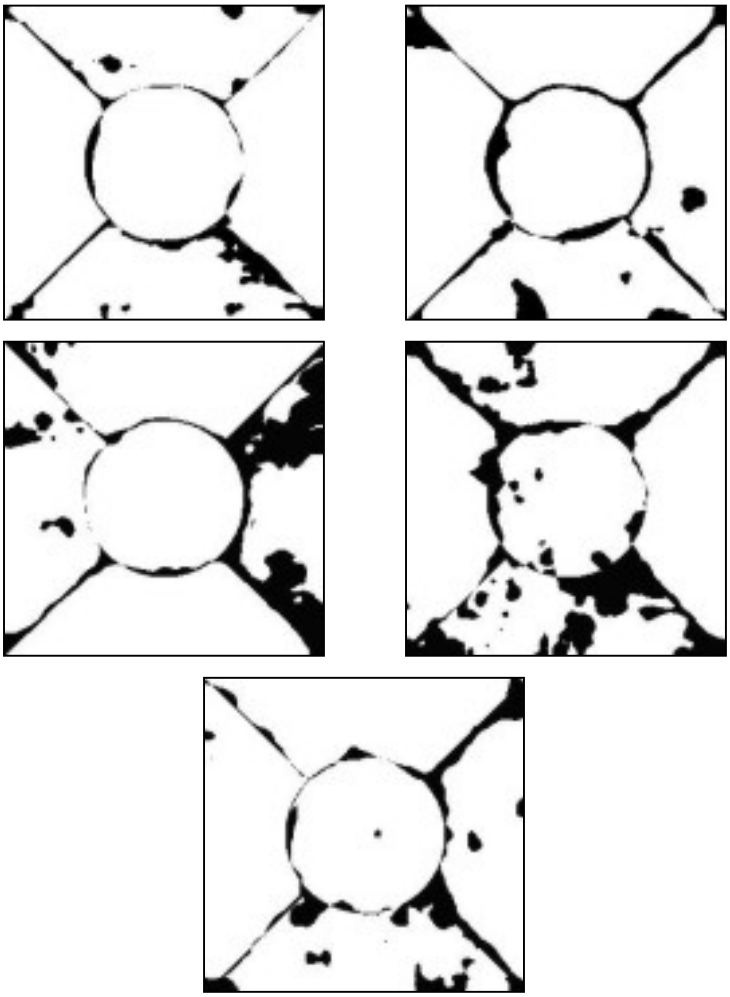

Figura 8 – Erro da segmentação dos Mosaicos da Figura 7.

REFERÊNCIAS BIBLIOGRÁFICAS

BRODATZ, P., **A photographic Album for Artists & Designers**, New York, 1966, Dover Publications, Inc.

CAMPBELL, F. W. and Robson, J. G., **Application of Fourier Analysis to the Visibility of Gratings**, J. PHYSIOLOGY, 197, p. 551-566, 1968

CONNERS, R. W., Harlow, C. A. **A theoretical comparison of texture algorithms**. IEEE TRANS. PATTERN ANAL. AND MACHINE INTELL., 2(3), 1980, p. 204-221.

DE VALOIS, R. L., Albrecht, D. G. and Thorell, L. G., **Spatial-frequency selectivity of cells in macaque visual cortex**, VISION RES., 22, p. 545-559, 1982

du BUF, J. M. H., Kardan, M., Spann, M. **Texture Feature Performance for Image Segmentation**. PATTERN RECOGNITION, 23(3), p. 291-309, 1990.

DUDA, R. O. and Hart, P. E., **Pattern Classification and Scene Analysis**, New York, 1973, Jonh Wiley & Sons.

DUNN D. F., Higgins, W. E., Wakeley, J., **Texture Segmentation Using 2-D Gabor Elementary Functions**. IEEE TRANS. PATTERN ANAL. AND MACHINE INTELL., 22(2), p. 130-149, 1994.

DUNN, D. F., Higgins, W. E., **Optimal Gabor Filters for Texture Segmentation**. IEEE TRANS. IMAGE PROCESSING, 4(7), p. 947-964, 1995.

GROSSMAN, A., and Morlet, J., **Decomposition of Hardy Functions into Square Integrable *Wavelets* of Constant Shape**, SIAM J. MATH., 15, p. 723-736, 1984

HARALICK, R. M., Shanmugam, K. and Dinstein, I.**, Textural Features for Image Classification**, IEEE TRANS. On SMC, 3, p. 610-621, 1973.

JULEZ, B., **Experiments in the Visual Perception of texture**, SCIENTIFIC AMERICAN, 212, 34-43, 1975

Julez, B., **Visual Pattern Discrimination**, IRE TRANSACTIONS INF. THEORY, 8, 84-92, 1962.

LAWS, K. I., **Rapid Texture Identification**. PROC. OF THE SPIE CONFERENCE ON IMAGE PROCESSING FOR MISSILE GUIDANCE, p. 376-380, 1980.

MALLAT, S.G., **A Wavelet Tour of Signal Processing**, San Diego, 1998, Academic Press, 577 p.

MALLAT, S., Zhong, S. **Characterization of Signals from Multiscale Edges**, IEEE TRANS. ON PAMI, 14(7), p. 710-732, 1992.

MALLAT, S., Zhong, S., **Wavelet Maxima Representation**. *WAVELETS* AND APPLICATIONS-PROCEEDINGS OF INT. CONF., Marseille(1989), 1992, Yves Meyer Ed., Masson.

OJALA, T., Pietikäinen, M., Harwood, D. **A comparative Study of Texture Measures with Classification Based on Feature Distributions**. PATTERN RECOGNITION, 29(1), p. 51-59, 1996.

OJALA, T., Valkealahti, K., Oja, K., Pitikäinen, M., **Texture Discrimination With Multidimensional Distributions of Signed Gray-level Differences.** PATTERN RECOGNITION, 34, p 727-739, 2001.

PAGAMISSE, A., Dutra L.V. **Segmentação de Textura pela Transformada Wavelet Diádica.** PESQUISAS EM CIÊNCIAS GEODÉSICAS, 2, p. 152-174, Curitiba, PR, 2002.

RANDEN, J., Husøy, H., **Filtering for Texture Classification: A Comparative Study.** IEEE TRANSACTIONS ON PAMI, 21(4), p. 291-310, 1999.

STRAND, J., Taxt, T. **Local Frequency Features for Texture Classification**. PATTERN RECOGNITION, 27(10), p. 1397-1406, 1994.

TUCERYAN, M. and Jain, A. K. **Texture Analysis**, in HANDBOOK OF PATTERN RECOGNITION AND COMPUTER VISION, World Scientific Publishing Company, p. 610-621, 1973.

WELDON, T. P., Higgins, W. E. and Dunn, D. F., **Efficient Gabor Filter Design for Texture Segmentation**. PATTERN RECOGNITION, 29(12), p. 2005-2015, 1996.

WESZKA, J. S., Dyer, C. R., Rosenfeld, A. **A comparative study of texture measures for terrain classification.** IEEE TRANS. SYST., MAN, CYB., 6(4), p. 269-285, 1976.