

SEGMENTAÇÃO DE TEXTURAS PELA TRANSFORMADA WAVELET DIÁDICA

AYLTON PAGAMISSE⁽¹⁾
LUCIANO VIEIRA DUTRA⁽²⁾

⁽¹⁾Universidade Estadual Paulista - Unesp
Departamento de Matemática, 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 - Em imagens de sensoriamento remoto faz-se necessário, em muitos casos, a discriminação de texturas. Intuitivamente, textura está relacionada a variações tonais na imagem, sendo propriedade de uma região e não de um ponto. Essas variações podem ser vistas como descontinuidades locais. Em imagens, descontinuidades estão, em geral, associadas a bordas que geralmente são usadas para descrever a informação estrutural dos contornos de objetos. Detecção de bordas é geralmente visto como um processo que reduz a quantidade de dados representando um sinal mas não preservando toda a informação. Mallat e Zhong mostraram que bordas em multiescala da transformada wavelet diádica fornecem uma representação completa e estável do sinal. A partir dessa representação a discriminação das texturas é obtida ao se quantificar o conteúdo de bordas detectado pela transformada wavelet. Como textura é um fenômeno local e não pontual, essa quantificação é obtida homogeneizando-se as saídas da transformada wavelet e medindo-se a quantidade de energia local, além disso, para efeito de segmentação busca-se em cada nível da transformada uma banda de máxima separabilidade. Todo esse procedimento é realizado em amostras representativas das texturas e os parâmetros obtidos são utilizados na imagem a ser segmentada implicando em baixo custo computacional.

ABSTRACT - In remote sensing images it becomes necessary, in many cases, the discrimination of textures. Intuitively, texture is related to tonal variations in the image, being property of a region and not of a point. These variations can be seen as local discontinuities. In images, discontinuities are, in general, associated to edges that generally are used to describe the structural information of the object contours. Edge detection generally is seen as a process that reduces the amount of data representing a signal but not preserving all the information. Mallat and Zhong had shown that edges in multiscale of dyadic wavelet transform supply a complete and steady representation of the signal. From this representation the discrimination of the textures is obtained by quantifying the content of edges detected by wavelet transform. As texture is a local phenomenon and not punctual, this quantification is obtained homogenizing the wavelet transform outputs and measuring its local amount of energy, moreover, for segmentation effect a band of maximum separability in each level of the transform is sought. All this procedure is carried through in representative samples of the textures and the obtained parameters are used in the image to be segmented, implying in low computational cost.

1 INTRODUÇÃO

Dentre as várias características das imagens 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. Devido a essa diversidade torna-se impossível estabelecer uma 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. Além disso, ela exerce um importante papel no sistema

visual humano, fornecendo informação para reconhecimento e interpretação.

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).

A transformada wavelet diádica apresentada a seguir é uma representação invariante por translação e se caracteriza por decompor um sinal num esquema multiresolucional, preservando sua energia, (Mallat e Zhong, (1992)(1)(2)). Essa representação de um sinal é redundante, no entanto, a fim de se realizar a discriminação de texturas, é interessante que se tenha uma representação invariante por translação. Quanto à redundância, cabe salientar que cada um dos canais gerados pela decomposição traz alguma informação adicional, muitas vezes não presente em outros canais e isso será explorado dentro de um contexto adequado.

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,

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

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

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

Se existirem $A_1, B_1 > 0$ tais que

$$\forall \omega \in \mathbb{R}, A_1 \leq \sum_{j=-\infty}^{+\infty} |\hat{\psi}(2^j \omega)|^2 \leq B_1$$

então fica assegurado que todo o eixo de frequência é coberto por dilatações de $\hat{\psi}(\omega)$ por $[2^j]_{j \in \mathbb{Z}}$ (Mallat, 1998), de modo que $\hat{f}(\omega)$ e $f(x)$ podem ser recuperadas a partir de sua WT diádica. A wavelet de reconstrução $\chi(x)$ é qualquer função cuja transformada de Fourier satisfaça

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

e a função $f(x)$ é recuperada a partir de sua WT diádica com a soma

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

2.1 Uma Família de Wavelets Diádicas

A fim de poder computar a transformada wavelet diádica discreta por um banco de filtros, com um algoritmo rápido, deve-se construir a wavelet de modo apropriado, satisfazendo algumas relações de escala diádicas. 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 $\hat{H}(0) = 1$. Como nos casos de bases wavelets ortogonais e biortogonais (Mallat, 1998), construímos uma função de escala cuja transformada de Fourier é dada por,

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

Vamos supor que $\phi \in L^2(\mathbb{R})$. A wavelet correspondente ψ tem uma transformada de Fourier dada por

$$\hat{\psi}(\omega) = \hat{G}\left(\frac{\omega}{2}\right) \hat{\phi}\left(\frac{\omega}{2}\right) \quad (3)$$

Note que ϕ e ψ possuem suporte compacto pois h e g tem um número finito de coeficientes não nulos.

Uma família de funções satisfazendo as relações de escala acima pode ser construída a partir de uma B-splines de ordem n obtida pela 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 $(0,1]$. O suporte da B-spline de ordem n é o intervalo $[0, n]$ e sua transformada de Fourier é dada por

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

Se $\phi = B^n$, então a relação de escala (4.3) diz que

$$\hat{H}(\omega) = \frac{\hat{B}^n(2\omega)}{\hat{B}^n(\omega)} = e^{-\frac{ik\omega}{2}} \left[\cos \frac{\omega}{2} \right]^{n+1} \quad (4)$$

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

$$h_k = \begin{cases} \frac{1}{2^{n+1}} \binom{n+1}{\frac{n+1}{2} + k} & ; |k| \leq n+1 \\ 0 & \text{caso contrário} \end{cases} \quad (5)$$

para n ímpar

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

$$h_k = \begin{cases} \frac{1}{2^{n+1}} \binom{n+1}{\frac{n}{2} + 1 + k} & ; -\frac{n}{2} - 1 \leq k \leq \frac{n}{2} \\ 0 & \text{caso contrário} \end{cases} \quad (6)$$

para n 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 $n = \text{grau da B-spline}$,

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

Queremos que ψ^n seja a derivada primeira de uma B-spline. Para que isto ocorra, tomamos

$$\hat{G}(\omega) = -ie^{-\frac{i\omega}{2}} \sin \frac{\omega}{2}. \quad (7)$$

A transformada de Fourier da wavelet resultante será

$$\hat{\psi}^n(\omega) = \hat{G}(\frac{\omega}{2}) \hat{\phi}(\frac{\omega}{2}) = -i \frac{\omega}{4} e^{-\frac{i\omega(1+k)}{4}} \left(\text{sinc} \frac{\omega}{4} \right)^{n+2} \quad (8)$$

donde ψ^n é a derivada primeira de uma B-spline de grau $n+1$ centrada em $t = (1+k)/4$.

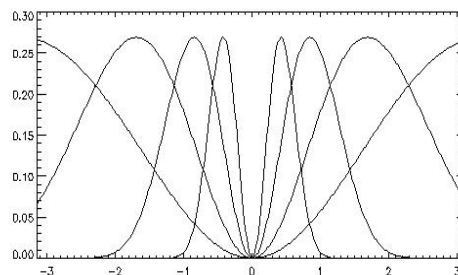


Figura 1 - Gráfico do módulo da equação (4.7) para $0 \leq n \leq 3$.

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 $k \in \mathbb{Z}$.

2.2 Implementação no Caso Discreto

Para o modelo discreto, assim como o modelo contínuo acima, a WT diádica de um sinal em $\ell^2(\mathbb{Z})$, isto é, sinais discretos de energia finita, é implementada como o operador linear limitado $W: \ell^2(\mathbb{Z}) \rightarrow \ell^2(I)$, onde, $I = \{1, 2, \dots, J+1\} \times \mathbb{Z}$, consistindo de $J+1$ operadores lineares

$$W_j: \ell^2(\mathbb{Z}) \rightarrow \ell^2(\mathbb{Z}), j = 1, 2, \dots, J+1.$$

Os operadores W_j são operadores de convolução com filtros cujas respostas à impulso são dados 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 escala 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 $H = \{h_k\}$ e $G = \{g_k\}$. Então as relações de escala (2) e (3) de $\hat{\phi}(\omega)$ e $\hat{\psi}(\omega)$ nos dizem que

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

donde, definindo $S_0 f = f$, 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

$$W_j f = W_{2^j} f, S_j f = S_{2^j} f, 0 \leq j \leq J \text{ e}$$

(*) denota convolução.

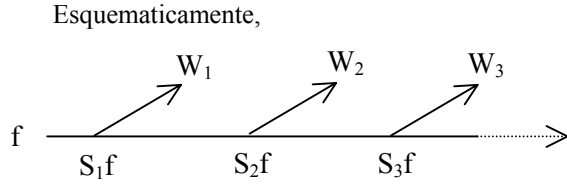


Figura 2 - Esquema de decomposição 1D pela transformada wavelet diádica.

Se o sinal original $\{S_0 f(n); n \in \mathbb{Z}\}$ tem N amostras, podemos computar no máximo $\log_2(N)$ escalas da transformada wavelet diádica discreta.

2.3 Transformada Wavelet Diádica de Imagens

O modelo bi-dimensional é construído a partir do modelo um-dimensional através do produto tensorial. As mesmas propriedades são válidas, e a implementação discreta é obtida por meio de convoluções por filtros unidimensionais nas linhas e colunas do sinal discreto da seguinte forma.

Considere as wavelets $\psi^1(x, y)$ e $\psi^2(x, y)$, correspondendo às direções horizontais e verticais, dadas por,

$$\psi^1(x, y) = \psi(x)2\phi(y) \text{ e } \psi^2(x, y) = 2\phi(x)\psi(y)$$

Se

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

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

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

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

$$\forall (\omega_x, \omega_y) \in \mathbb{R}^2 - \{(0, 0)\},$$

$$A_2 \leq \sum_{k=1}^2 \sum_{j=-\infty}^{+\infty} \left| \hat{\psi}^k(2^j \omega_x, 2^j \omega_y) \right|^2 \leq B_2$$

Então existem wavelets de reconstrução $\{\chi^k(x, y)\}_{1 \leq k \leq 2}$ tais que

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

donde

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

As wavelets acima possuem as relações de escala

$$\hat{\psi}^1(\omega_x, \omega_y) = \hat{\psi}(\omega_x) \hat{\phi}\left(\frac{\omega_y}{2}\right) = G\left(\frac{\omega_x}{2}\right) \hat{\phi}\left(\frac{\omega_x}{2}\right) \hat{\phi}\left(\frac{\omega_y}{2}\right) \quad (9)$$

$$\hat{\psi}^2(\omega_x, \omega_y) = \hat{\phi}\left(\frac{\omega_x}{2}\right) \hat{\psi}(\omega_y) = \hat{\phi}\left(\frac{\omega_x}{2}\right) G\left(\frac{\omega_y}{2}\right) \hat{\phi}\left(\frac{\omega_y}{2}\right) \quad (10)$$

As relações de escala em (9) e (10), permitem-nos uma implementação rápida da transformada wavelet diádica. Para isto, seja H o filtro discreto dado por $H = \{h_k\}$ e G o filtro discreto dado por $G = \{g_k\}$ da mesma forma que no caso 1D, H_p e G_p são os filtros discretos obtidos ao colocar-se $2^p - 1$ zeros entre cada coeficiente dos filtros H e G . Desse modo, a decomposição diádica discreta de um sinal bidimensional f é feita da seguinte forma

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

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

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

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.

Esquemáticamente,

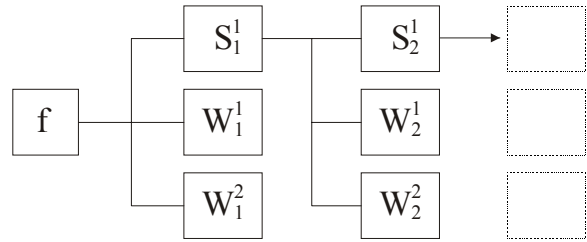


Figura 3 - Esquema de Decomposição 2D pela transformada wavelet diádica.

onde, em S temos as versões suavizadas do sinal que são utilizadas na próxima escala, enquanto que nos W 's temos os detalhes que foram extraídos do sinal nas direções horizontal e vertical.

3 MODELO EXPERIMENTAL

Nossa asserção básica é que medidas da distribuição da energia local, resultante da decomposição do sinal dada pela transformada wavelet diádica discreta nas suas várias resoluções são capazes de discriminar inúmeras texturas. Isto decorre do fato que $|W^k f(u, v, 2^j)|^2$ é maior quando 2^j corresponde à escala das estruturas componentes das texturas.

O uso da energia local é bastante utilizado não só no contexto da transformada wavelet mas também no da transformada de Gabor ou mesmo nos de matriz de coocorrência (Connors e Harlow, 1980) (Dunn et al, 1994) (Dunn e Higgins, 1995) (Haralick et al, 1973).

O procedimento a fim de realizar as medidas da distribuição da energia local, no nível J da transformada wavelet diádica, consiste nas seguintes tarefas:

- decomposição wavelet diádica
- tomada do valor absoluto das saídas da decomposição

- homogeneização
- cálculo das energias locais.

Na primeira etapa, onde é feita a decomposição wavelet, as informações correspondentes a diferenças entre regiões de texturas estão concentradas em diferenças nas variâncias dos atributos iniciais, pois 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 médias. A homogeneização subsequente 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, ela 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. Alguns autores, (Randen, 1997) (Weldon et al, 1996), 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 (Gomes e Velho, 1994) é a que se utilizou aqui.

Após a homogeneização, obtemos duas bandas por nível da transformada wavelet. Cada banda traz consigo alguma informação sobre a imagem, visto que são oriundas de orientações específicas. A fim de se obter a maior separabilidade possível entre as texturas, fazemos a 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, temos em mãos três bandas. Devemos então selecionar qual a melhor delas. Para isso, utilizamos 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. De modo geral, a banda dada pela composição das direções canônicas possui o maior coeficiente. A energia local é então tomada em regiões de mesmo tamanho que as amostras e consiste de uma média. Como esse procedimento é feito em amostras das texturas, tem baixo custo computacional.

3.1 Realização Experimental

Para cada par de texturas T_1 e T_2 a serem segmentadas, realizamos o seguinte procedimento:

1 - Tomamos amostras representativas de cada textura, e para cada nível da transformada wavelet diádica, obtemos 2 bandas, b_1 , b_2 , conforme os filtros

gerados pela implementação discreta da transformada wavelet diádica.

2 - A partir dessas bandas, obtemos a banda com a máxima separabilidade entre as duas texturas.

3 - Comparamos a discriminabilidade 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, conforme a figura 4.

No caso de mais que duas texturas, aplica-se o mesmo procedimento para cada par de texturas e usa-se alguma técnica de segmentação. No nosso caso, utilizamos o seguinte algoritmo combinatorial, que além de ser facilmente implementado, apresentou bons resultados.

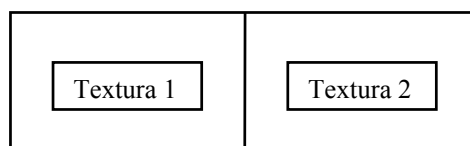


Figura 4 - Modelo de par de texturas.

3.2 Algoritmo Combinatorial

Para cada par de texturas, geramos uma banda que discrimina maximamente este par, num treinamento com amostras dessas duas texturas. Aplicando o resultado desse treinamento numa imagem composta por n texturas, obteremos $\frac{n(n-1)}{2}$ bandas. Assim, para cada pixel da

imagem obteremos 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 a textura 1 ou a textura 2, na banda 1. Na posição 2, o vetor terá o valor 1 ou 3, caso o pixel tenha sido classificado como pertencendo a 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 a textura $n-1$ ou a textura n , na banda $n(n-1)/2$. Por exemplo, consideremos um pixel pertencendo a textura 1, num mosaico de 4 texturas. Tipicamente, seu vetor atributo será da forma, $(1,1,1,x,y,z)$. Temos 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 esses parâmetros, gera-se a banda ótima, 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.

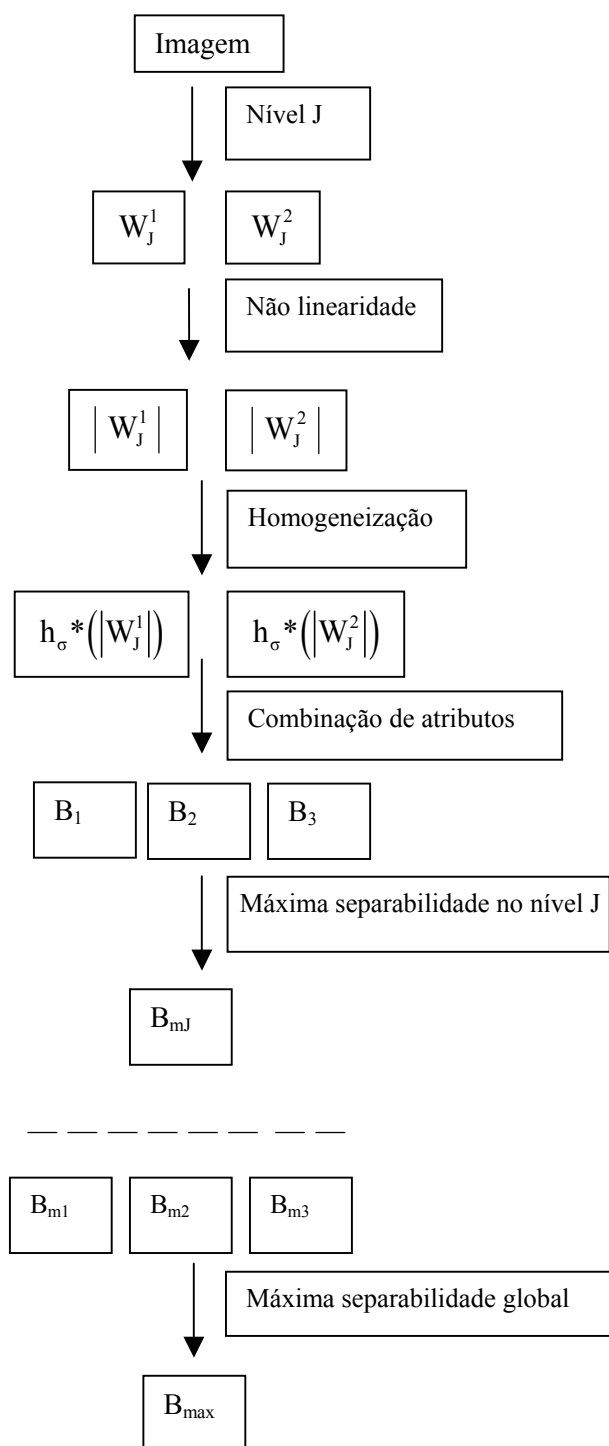


Figura 5 - Esquema representativo do modelo experimental

4 APLICAÇÕES DO MODELO EXPERIMENTAL

Vamos aplicar o modelo experimental um conjunto de pares de texturas e a mosaicos formados por cinco texturas e comparar os resultados obtidos com vários métodos existentes, apresentando os resultados dispostos

em tabelas comparativas. De antemão, a partir das médias apresentadas nas tabelas comparativas, observamos que, no caso de pares de texturas, nosso método mostrou-se, na média dos três pares, superior a todos os outros, enquanto que em mosaicos com 5 texturas sua performance também foi eficiente.

A maior parte das texturas utilizadas na comparação de modelos de discriminação de texturas, são extraídas do álbum “A Photographic 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 a regra, optamos por utilizar algumas dessas texturas a fim de avaliar a performance do método apresentado e compará-lo a resultados encontrados na literatura.

A comparação dos resultados aqui obtidos 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, havia poucas comparações entre esquemas de extração de atributos de textura. Podemos citar: Weszka et al (1976), Connors 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 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. Na tabela 2 apresentamos alguns dos resultados de Randen e Husoy juntamente com aqueles obtidos pelo nosso método.

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 6 são utilizados em Randen e Husoy em sua extensiva avaliação e os resultados obtidos

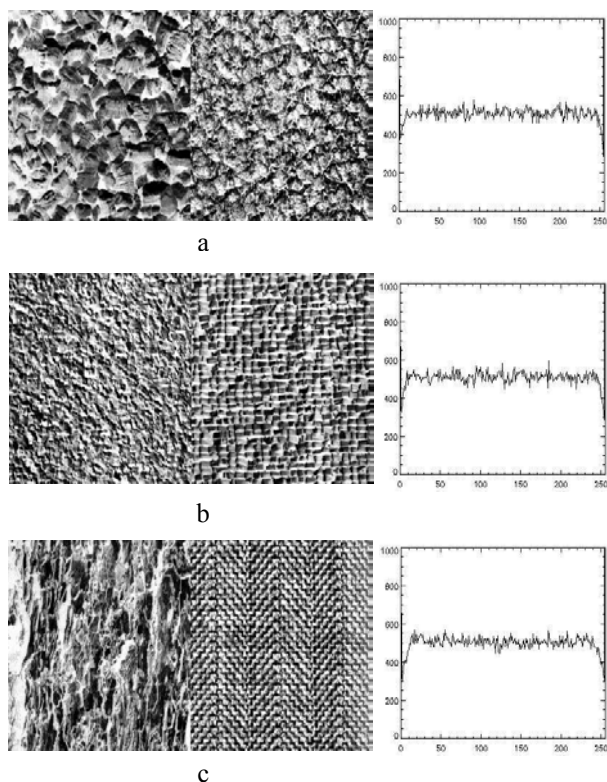


Figura 6 - (a) par de texturas D05D92 e histograma, (b) par de texturas D04D84 e histograma e (c) par de texturas D12D17 e histograma.

pelos inúmeros métodos por eles avaliados, encontram-se apresentados na tabela 2.

A fim de que as imagens da figura 6 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, podemos observar como os histogramas na figura 6 são praticamente idênticos e unimodais.

Na figura 7, apresentamos a banda de máxima separabilidade do par de texturas D005D092 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 threshold. No nosso caso, o threshold foi feito a partir da energia obtida em cada amostra, ou seja, de modo automático. Os demais pares de texturas, apresentam um comportamento semelhante, não havendo necessidade de apresentá-los.

Os passos utilizados para a geração das bandas de máxima separabilidade são aqueles descritos na seção 3 quando da descrição do modelo experimental. Utilizamos amostras 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.

No caso dos pares apresentados na figura 6, 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 e formadas por:

- figura 6(a): formada pelas texturas D005 e D092
- figura 6(b): formada pelas texturas D004 e D084
- figura 6(c): formada pelas texturas D012 e D017

As texturas em 6(a) caracterizam-se principalmente por diferenças na escala, enquanto que as texturas em 6(b) caracterizam-se pela repetitividade e nas texturas em 6(c) encontramos diferenças em orientação, escala e repetitividade.

Tabela 2 - Porcentagem de erro para métodos de segmentação aplicados aos pares de texturas D004D084, D005D092, D012D017. Resultados extraídos de Randen e Husoy.

Método	D04D84	D05D92	D12D17
Filtros de Law	24,8 %	8,6 %	3,3 %
Banco de Filtros Diádicos de Gabor	3,6 %	11,3 %	2,0 %
Gabor arbitrário	6,5 %	15,6 %	1,2 %
DCT	6,4 %	2,5 %	2,2 %
Wavelet Frame	4,4 %	4,7 %	0,4 %
QMF	6,7 %	6,0 %	0,8 %
Coocorrência	1,9 %	3,3 %	4,8 %
AR	11,3 %	3,0 %	1,9 %
Gabor full rate	6,5 %	15,6 %	1,2 %
Daubechies 4 full rate	5,7 %	8,2 %	0,6 %
Daubechies 6 full rate	4,7 %	6,1 %	0,6 %
Daubechies 8 full rate	4,9 %	4,7 %	1,3 %
Daubechies 10 full rate	4,4 %	5,1 %	1,2 %
F 16 b, full rate	8,1 %	8,2 %	0,8 %
F 2 1 smpl, full rate	7,2 %	7,4 %	2,3 %
F 2 1 09, full rate	6,7 %	6,0 %	2,1 %
Mahalanobis-Singh	1,3 %	1,4 %	24,6 %
Fisher, closed form	1,4 %	7,4 %	4,6 %
Fisher, gradient	1,9 %	1,6 %	3,0 %
Dunn	2,8 %	1,0 %	25,4 %
Eigenfilter	4,0	4,7	3,7
Opt. Repr. Filter Gabor	5,5	36,2	5,8
JMS	0,7	28,6	8,3
JU	0,7	5,1	2,1
JF	0,7	5,1	2,1
Optimal Gabor filter (best)	5,3	10,8	3,3
Prediction error filter	0,9	31,7	4,6
BackProp. NN mask size=11	32,1	43,6	30,0
BackProp. NN mask size=21	30,8	48,6	28,8
4,6,10 filter opt. Gabor filter bank (melhor)	3,8	12,8	3,8
Wavelets Diádicas	0,8 %	2,6 %	2,4 %

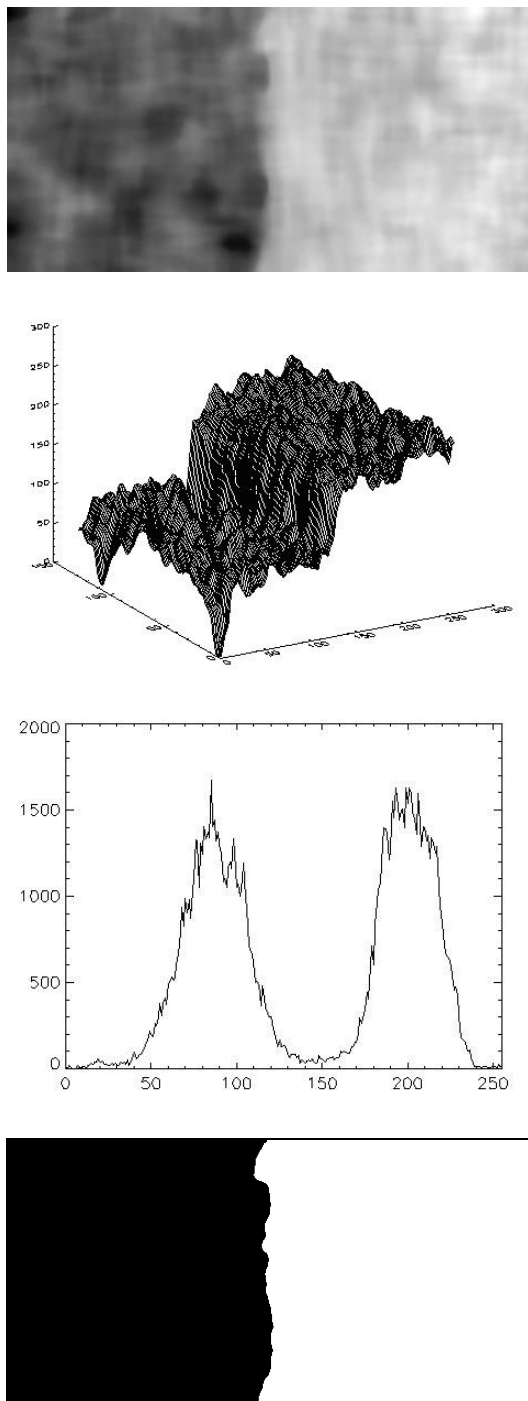


Figura 7 – De cima para baixo: banda de máxima separabilidade para o par de texturas D005D092, sua superfície de tons de cinza, seu histograma e resultado da segmentação.

4.2 Discussão dos Resultados

Apesar do método aqui apresentado não ser o de melhor performance quando aplicado às três texturas, a diferença, em todos os casos, é muito pequena em relação aos métodos que apresentam os melhores resultados. Quanto à boa performance de JMS, JU e JF no par D004D084, 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.

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 7. 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 é muito superior a maioria dos métodos descritos na tabela 2 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 bastante robusto.

4.3 Segmentação de Mosaicos de Cinco Texturas

Seguindo a apresentação de Randen e Husoy com relação ao tipo de testes efetuados, vamos aplicar 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 8 abaixo apresenta esses mosaicos, enquanto que a tabela 3 mostra a performance de vários métodos utilizados para segmentá-los, conforme Randen e Husoy, juntamente com os resultados por nós obtidos. As texturas desses mosaicos não são oriundas exclusivamente de Brodatz sendo que o mosaico Nat-5v é formado por texturas de MIT Vision Texture Database (<http://whitechapel.media.mit.edu/vismod/imagery/VisionTexture/vistex.html>).

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 mosaicos 5c e 5v, conforme designação de Randen e Husoy, são formados pelas seguintes texturas:

- Mosaico 5c: D077, D084, D055, D053 e D024, (Brodatz.,1966)

- Mosaico 5v: Fabric.0000, Fabric.0017, Flowers.0002, Leaves.0006 e Leaves.0013, (Mit).

Os resultados obtidos por nosso modelo experimental, juntamente com resultados dos outros métodos, são mostrados na tabela 3.

Tabela 3 – Porcentagem de erro de erro para métodos de segmentação aplicados aos mosaicos Nat-5c e Nat-5v. Resultados extraídos de Randen e Husoy.

Método	5 c	5 v	Média
Filtros de Law	9,7 %	25,7 %	18,7%
Banco de Filtros Diádicos de Gabor	10,7%	34,8 %	22,7%
Banco de filtros de Gabor	8,2%	34,0 %	21,1%
DCT	13,2%	27,0 %	15,1%
Coocorrência	9,9 %	27,0 %	18,4%
AR	19,6 %	19,4 %	19,5%
Daubechies 4	8,7 %	22,8 %	15,7%
F 16 b	8,7 %	18,9 %	13,8%
Eigenfilter	12,1 %	27,0%	19,5%
Opt. Repr. Gabor filter bankk.	7,2%	60,5%	33,9%
Prediction error filter	14,7%	31,1%	22,9%
4-filter opt. Gabor filter bank.	24,7%	54,5%	39,6%
6-filter opt. Gabor filter bk.	16,3%	34,5%	25,4%
10-filter opt. Gabor filt. bk.	16,1%	29,4%	22,7%
JMS	16,9%	36,3%	26,6%
JU	12,7%	33,0%	22,8%
BackProp. NN mask size=11	47,4%	65,0%	56,2%
BackProp. NN mask size=21	72,4%	68,7%	70,5%
Wavelets Diádicas	11,7%	17,7%	14,7%

4.4 Discussão dos Resultados

Podemos observar a boa performance de nosso método quando comparado à maioria dos outros e em termos de média, para esses dois mosaicos, ele é inferior apenas ao método F 16 b. 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. As figuras 8 e 9 apresentam os resultados da segmentação dos mosaicos com cinco texturas assim como 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 no mosaico 5v encontramos uma variabilidade bem maior. Isso se deve a natureza homogênea das texturas em 5c e a não homogeneidade observada no mosaico 5v. Nesse mosaico, encontramos regiões que pode-se dizer não pertencerem à textura na qual estão contidas. É como se

fossem “buracos” dentro dessas texturas. 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.



Figura 8 – De cima para baixo: mosaico Nat-5c, resultado da segmentação e imagem erro da segmentação.

4.5 Observações

O método apresentado nesse trabalho pode ser implementado com custo computacional muito baixo pois já no primeiro nível da transformada wavelet obtém-se bons resultados em muitos tipos de texturas. Cabe ao usuário fazer a escolha em função do tipo de texturas envolvidas.

Nesse método, não consideramos o uso de mais direções na construção da transformada wavelet. Esse é o objeto de um próximo trabalho. As bases para isso já estão postas nesse primeiro método.



Figura 9 – De cima para baixo: mosaico Nat-5v, resultado da segmentação e imagem erro da segmentação

REFERÊNCIAS

Brodatz, P., **A photographic album for artists & designers**, New York, 1966, Dover Publications, Inc.

Connors, R. W., Harlow, C. A. A theoretical comparison of texture algorithms. **IEEE Trans. Pattern Anal. And Machine Intell.**, V. 2(3), p. 204-221, 1980.

du Buf, J. M. H., Kardan, M., Spann, M. Texture feature performance for image segmentation. **Pattern Recognition**, V. 23(3), p. 291-309, 1990.

Duda, R. O. and Hart, P. E., **Pattern classification and scene analysis**, New York, 1973, John 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., V. 22(2), p. 130-149, 1994.

Dunn, D. F., Higgins, W. E., Optimal Gabor filters for texture segmentation. **IEEE Trans. Image Processing**, V. 4(7), p. 947-964, 1995.

Gomes, J. e Velho, L., **Computação Gráfica: Imagem**, Rio de Janeiro, 1994, IMPA-SBM, 421p.

Grossman, A., and Morlet, J., Decomposition of Hardy functions into square integrable wavelets of constant shape, **SIAM J. Math.**, vol. 15, p. 723-736, 1984

Haralick, R. M., Shanmugam, K. and Dinstein, I., Textural features for image classification, **IEEE Trans. On SMC**, V. 3, p. 610-621, 1973.

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, 577p.

Mallat, S., Zhong, S.⁽¹⁾. Characterization of Signals from Multiscale Edges, **IEEE Trans. on PAMI**, V. 14(7), p. 710-732, 1992.

Mallat, S., Zhong, S.⁽²⁾, 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**, V. 29(1), p. 51-59, 1996.

Randen, T., **Filter and filter bank design for image texture recognition**, 1997. 130p. PhD Thesis - , Norwegian Un. Of Science. And Technology, 1997.

Strand, J., Taxt, T. Local frequency features for texture classification. **Pattern Recognition**, V. 27(10), p. 1397-1406, 1994.

Weldon, T. P., Higgins, W. E. and Dunn, D. F., Efficient Gabor filter design for texture segmentation. **Pattern Recognition**, V. 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.**, V. 6(4), p. 269-285, 1976.

Randen, J., Husøy, H., Filtering for texture classification: A comparative study **IEEE Transactions on PAMI**, V. 21(4), p. 1999.