Two Texture Segmentation using M-Band Wavelet Transform

Mausumi Acharyya and Malay K. Kundu*
Machine Intelligence Unit, Indian Statistical Institute
203, B. T. Road, Calcutta - 700 035, INDIA
email: {res9522, malay}@isical.ac.in

Abstract

The M-band wavelet decomposition, which is a direct generalization of the standard 2-band wavelet decomposition has been applied to the problem of an unsupervised segmentation of two texture systems. Standard wavelets are not suitable for the analysis of high frequency signals with relatively narrow bandwidth. So in the present work we were motivated to use the decomposition scheme based on M-band wavelets, that yield improved segmentation accuracies. Unlike the standard wavelet decomposition which give a logarithmic frequency resolution, the M-band decomposition gives a mixture of a logarithmic and linear frequency resolution. Further motivation to use M-band wavelet filter for our texture analysis scheme is because, this decomposition yields a large number of subbands which is required for good quality segmentation.

Index - M-band wavelets, Texture segmentation, Feature extraction

1. Introduction

In this paper we focus on multichannel filtering approach to texture analysis. Successful application of multichannel filtering for texture segmentation is reported in [4] [3]. Various filtering techniques have been reported, but most popular channel filters that is in use are the Gabor filters . Randen *et. al.* [9] have examined the performance of multichannel segmentation schemes based on more general class of filters including Gabor filters. However a large combination of parameters makes texture discrimination using Gabor filters computationally expensive.

Successful application of wavelet theory to texture analysis have been reported using the multiresolution signal decomposition developed by Mallat [7]. The standard or the octave band wavelet decomposition imply fine frequency resolution in the low frequency compared to the high frequency region. The work of Chang and Kuo [3] indicates that the texture features are more prevalent in the intermediate frequency band. Laine *et. al.* [6] carried out studies on texture analysis based on this indication. They used multichannel wavelet frames for feature extraction. Representations obtained from both standard wavelets and wavelet packets were evaluated.

One of the drawback of standard wavelets is that they are not suitable for the analysis of high frequency signals with relatively narrow bandwidth. So in the present work we were motivated to use the decomposition scheme based on M-band wavelets, that yield improved segmentation accuracies. Unlike the standard wavelet decomposition which gives a logarithmic frequency resolution, the M-band decomposition gives a mixture of a logarithmic and linear frequency resolution. Moreover, since a large number of subbands is required in order to achieve a good quality segmentation, we were further motivated to use the M-band decomposition scheme because it generates a large number of subbands.

The proposed method of texture segmentation employs the filter response energy measures as features. A typical system setup consist of a filtering step followed by a local energy estimator. Basically the purpose of the filter is extraction of local frequencies, where one of the textures have low signal energy and the other texture have high energy. If this is accomplished a composite texture image can be segmented and the region classified by analysis of their energies. In multichannel approach one such filter is typically used for each channel.

In the present work we investigate the problem of segmentation of two texture systems by using a generalization of the wavelet transforms to the M-band case. We conjecture that the M-band wavelet transform has the potential to perform multiscale, multidirectional filtering of the images, since it is a tool to view signals at different scales and decomposes a signal by projecting it onto a family of functions generated from a single wavelet basis via its dilations and translations [1]. Various combinations of the M-band wavelet filter decompose the image at different scales and

*corresponding author

0-7695-0750-6/00 $10.00 © 2000 IEEE

orientations in the frequency plane. A local energy measure is applied to these bandpass sections which then give the texture features that can be classified with success.

Section 2 briefly overviews the M-band wavelet transform. Section 3 presents our methodology. Finally section 4 concludes with results and critical comments of the present work.

2. M-band wavelet transform

M-band wavelets are a direct generalization of the conventional wavelets. These M-band wavelets are able to zoom in onto narrowband high frequency components of a signal and they are reported to give better energy compaction than 2-band wavelets. M-band wavelets are a set of $M-1$ basis functions whose scaled and translated versions form a tight frame for the set of square integrable functions defined over the set of real numbers $L^2(R)$ [2]. For the $(M-1)$ wavelets, $\psi_i(x)$, $i = 1, \cdots, M-1$, given any function $f(x) \in L^2(R)$,

$$f(x) = \sum_{i=1}^{i=M-1} \sum_{j \in Z} \sum_{k \in Z} \langle f(x), \psi_{i,j,k}(x) \rangle \psi_{i,j,k}(x) \quad (1)$$

where Z represents the set of integers. The $\psi_{i,j,k}(x)$ functions are obtained by performing scaling and shifts to the corresponding wavelet $\psi_i(x)$,

$$\psi_{i,j,k}(x) = M^{j/2} \psi_i(M^j x - k) \quad (2)$$

The wavelet functions are defined from a unique, compactly supported scaling function $\psi_0(x) \in L^2(R)$ with support in $[0, \frac{(N-1)}{(M-1)}]$ by,

$$\psi_i(x) = \sqrt{M} \sum_{k=0}^{k=N-1} h_i(k) \psi_o(Mx - k). \quad (3)$$

The scaling function satisfies the equation,

$$\psi_0(x) = \sqrt{M} \sum_{k=0}^{k=N-1} h_0(k) \psi_0(Mx - k)$$

where h_0 is the unitary scaling vector of length $N = MK$ and is characterized by the constraints,

$$\sum_{k=0}^{k=N-1} h_0(k) = \sqrt{M} \text{ and } \sum_{k=0}^{k=N-1} h_0(k) h_0(k + Ml) = \delta l$$

The $(M-1)h_i$ wavelet vectors of length N satisfy,

$$\sum_{k=0}^{k=N-1} h_i(k) h_j(k + Ml) = \delta(l)\delta(i - j).$$

The scaling function and the $M-1$ wavelet functions also define a multiresolution analysis. A multiresolution analysis is a sequence of approximation spaces for $L^2(R)$. If the space spanned by the translates of $\psi_i(x)$ for fixed j and $k \in Z$ is represented by $W_{i,j}$, then it can be shown that,

$$W_{0,j} = \begin{matrix} i = M-1 \\ i = 0 \end{matrix} \oplus W_{i,j-1} \text{ and } \lim_{j \to \infty} W_{0,j} = L^2(R). \quad (4)$$

Thus the $W_{0,j}$ spaces form a multiresolution space for $L^2(R)$.

3. Computing and integrating texture features

3.1. M-band wavelet filters

The objective of the filtering and computation of the local energy measure, is to transform the edges between textures into detectable discontinuities. The filter bank in essence is a set of bandpass filters with frequency and orientation selective properties. An eight tap, M=4 band orthogonal and linear phase wavelet filter [1] is used to decompose the textured images into MXM-channels, by applying the M-band transform in the horizontal and vertical directions separably, but without downsampling giving overcomplete representations of the image. Various combinations of this filter decomposes the image into diferent scales and orientations in the frequency plane. M-band 2-D wavelet filters, are denoted by $\psi_{i,j}$, for i,j=1,2,3,4 with M=4. The i,j^{th} resolution cell is achieved by the filter $H_{i,j} = \psi_{i,j}\psi^*_{i,j}$, where $H_{i,j}$'s are the transfer functions corresponding to the filter impulse responses $\psi_{i,j}$'s, for i,j= 1,2,3,4 with M=4. Since the spectral response to edges of an image is strongest in direction perpendicular to the edge and decreases as the look direction of the filter approaches that of the edge. Therefore edge detection is envisaged by using 2-D filters that are lowpass along edge direction and highpass along the orthogonal direction. A typical edge detection filter corresponding to a particular direction covers a certain region in the 2-D spatial frequency domain. Based on this concept several wavelet decomposition filters are possible.

Since the filter system used is orthogonal and has quadrature mirror filter structure, *i.e.* $\sum_{i=1}^{M} \sum_{j=1}^{M} \psi_{i,j}\psi^*_{i,j} = 1$, the resulting 2-D filters treats all the frequencies in a resolution cell equally. The number of possible filter combinations depend on the number of bands (M). The decomposition filters for different directions in increasing resolutions are given as,

$\star$ Horizontal direction : $Filt_{hor1} = H_{12}$, $Filt_{hor2} = H_{12} + H_{13}$, $Filt_{hor3} = H_{12} + H_{13} + H_{14}$.

$\star$ Vertical direction : $Filt_{ver1} = H_{21}$, $Filt_{ver2} = H_{21} + H_{31}$, $Filt_{ver3} = H_{21} + H_{31} + H_{41}$.

Similarly the decomposition filters for diagonal ($Filt_{diag_j}$),

horizontal diagonal ($Filt_{hdiag_j}$) and vertical diagonal ($Filt_{vdiag_j}$), (where j=1,2,3) directions can be formed. These basically give a measure of texture energies along the different directions at different resolutions, the corresponding filtered images are given by h_{H_i}, h_{V_i} etc., for i=1,2,3.

3.2. Local energy estimator

The objective of the local energy extimator is to estimate the energy of the filter ouput in a local region. It is utilized for the purpose of transmitting areas in each channel where the bandpass frequency components are strong resulting in a high constant gray level and the areas where it is weak into a low constant gray level. We have used the most popular magnitude operation $|\cdot|$. The reason being, it is independent of the dynamic range of the input image and also of the filter amplification.

The nonlinear transform is succeeded by a Gaussian (smoothing) filter, $H_G(u,v) = \frac{1}{2\pi\sqrt{\sigma}} e^{\frac{1}{2\sigma^2}(u^2+v^2)}$. The feature image $F_k(x,y)$ corresponding to a filtered image $h_k(x,y)$ is expressed as,

$$F_k(x,y) = \frac{1}{W^2} \sum_{(a,b)\in W_{xy}} | \Psi(h_k(a,b)) | \qquad (5)$$

where, $k = H_i, V_i$ etc., $\Psi(\cdot)$ is the nonlinear function and W_{xy} is an WXW window centred at pixel with coordinates (x,y). The size W of the smoothing or the averaging window is an important parameter. More reliable measurement of texture feature calls for larger window sizes. On the other hand, more accurate localization of region boundaries calls for smaller windows. Gaussian weighted windows are preferrable over unweighted windows, because, averaging blurrs the boundaries between textured region and the former are likely to result in more accurate localization of texture boundaries.

Choice of σ of the smoothing filter : The choice of the space constant σ of the averaging filter is very crucial. Estimation of the local energy of an image with low spatial frequency requires the smoothing filter to have a wide unit impulse response, while narrower filters can be allowed for higher frequency content images.

In the present work we set the smoothing filter size based on the measure of the spectral content of the image. *Spectral flatness* is a measure of the overall image activity and is defined as the ratio of the arithmetic and the geometric mean of the Fourier coefficients [5]. For 2-D digital iamge this is expressed as,

$$SFM = \frac{\frac{1}{MN}\sum_{i=0}^{M-1}\sum_{j=0}^{N-1} | \hat{F}(i,j) |^2}{[\prod_{i=0}^{M-1}\prod_{j=0}^{N-1} | \hat{F}(i,j) |^2]^{\frac{1}{MN}}} \qquad (6)$$

$\hat{F}(i,j)$ is the $(i,j)^{th}$ Fourier coefficients. SFM has a dynamic range of $[0,1]$. Highly active image means SFM close to 1, that is the image have predominantly high frequencies, and requires smaller window, an image with low SFM have low spectral content and would require larger window while an image with moderate activity would require a moderate window size for smoothing.

We have experimentally found the values of σ for these three categories of image actitvities to lie within a range of $1 \leq \sigma \leq 8$, having a spatial extent of 11X11 to 31x31. Thus the size of the averaging window can be adaptively selected depending on the spectral content of the images. Since we do not use any fixed windowing operation our scheme can accomodate diverse set of textured images as input. This step gives the feature images.

3.3. Integrating the feature images

Clustering and postocessing: Segmentation algorithm accept as input a set of features and put a consistent labelling for each pixel. Fundamentally this can be considered a multidimensional data clustering problem. We emphasize on the feature extraction (representation) part in this work. So we have used a traditional unsupervised K-means clustering algorithm. The simple K-means clustering algorithm labels each pixel independently and do not take into account the high correlation between neighboring pixels. A more sophisticated algorithm should incorporate some neighborhood constraint into the segmentation process, such as relaxation labelling. So we have used median filtering to simulate the benefit of a local constraint.

4. Experimental results

To demonstrate the performance of our algorithm we have applied our texture segmentation algorithm to several two textured systems. Out of the 16 features possible for M=4, we considered only 13. The number of features could even be reduced in many textured systems.

In fig 1 for the texture pair $Nat3$ only 5 features were sufficient for successful segmentation of the images. We have given the results obtained by [10] and [8] for a comparative study.

To prove the efficacy of our algorithm, we have tested it over texture pairs that were visually not very distinct, but inspite of that they were well discriminated by our approach fig 2.

In fig 3 we give the experimental results of a texture pair. Figure also shows the texture features that we have used for obtaining the class map. The graphs in figure show the average features per column of the feature images, corresponding to FH_2, FD_2 and FHD_2 respectively, the figure also gives the texture feature images. It is evident from the

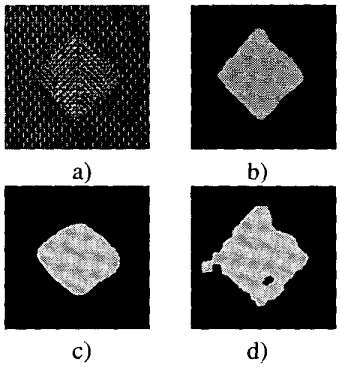

Figure 1. a) $Nat3$ and corresponding class maps using b) our method c) [10] d) [8]

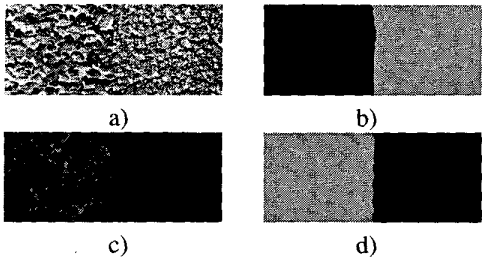

Figure 2. a) $D5D92$ b) corresponding class map c) $D9D24$ d) corresponding class map

graphs and the feature images, that the two textured regions have quite discernible features and good segmentation result is achieved in this case also.

Percentage of correct classification has been taken as a performance measure in this work. Percent of classification is 99.0% for $D5D92$, 99.5% for $D9D24$ and 99.6% for $D17D77$.

We have presented a multichannel filtering technique using M-band wavelets for texture segmentation. The filtering and the feature extraction operations account for most of the required computation, however our method is very simple and computationally less expensive and efficient. We have experimentally found that 3 to 5 features out of the 13 features are sufficient for good quality segmentation. So dimensionality of the feature space is greatly reduced. Also since the image can be decomposed into a large number of subbands we get a good representation of the image interms of frequencies in several directions at different resolutions. We therefore conjecture that this representation improves the quality of segmentation. The use of overcomplete wavelet representation of the textured images alleviates the problem of inaccurate texture boundary localization.

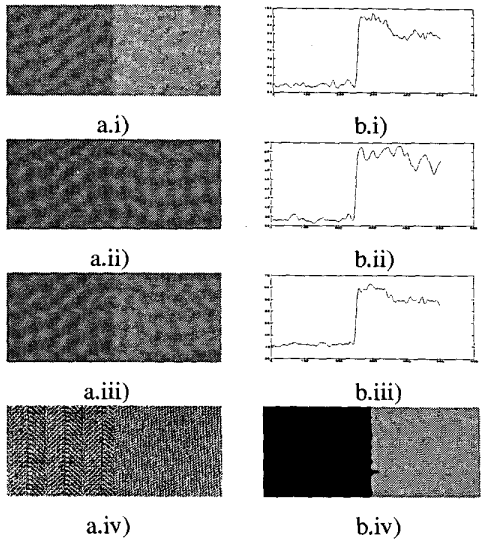

Figure 3. Feature images a.i) F_{H_2} a.ii) F_{D_2} a.iii) F_{HD_2} a.iv) $D17D77$ b.i)-b.iii) features averaged along columns b.iv) class map of D17D77 texture pair

References

[1] O. Atkin and H. Caglar. Design of efficient m-band coders with linear phase and perfect reconstruction properties. *IEEE Trans. Signal Processing*, 43(7):1579–1590, 1995.

[2] C. S. Burrus, A. Gopinath, and H. Guo. *Introduction to Wavelets and Wavelet Transform. A primer.* Prentice Hall International Editions, 1998.

[3] T. Chang and C. C. J. Kuo. Texture analysis and classification with tree structured wavelet transform. *IEEE Transactions on Image Processing*, 2(4):42–44, 1993.

[4] F.Farrokhnia and A.K.Jain. A multichannel filtering approach to texture segmentation. In *Proc. Computer Vision and Pattern Recognition*, pages 364–370, 1991.

[5] N. Jayant and P. Noll. *Digital Coding of Waveforms : Principles and Applications to Speech and Video.* Prentice Hall Publications, Englewood Cliffs, New Jersy, 1984.

[6] A. Laine and J. Fian. Frame representation for texture segmentation. *IEEE Trans. Image Process.*, 5(5):771–779, 1996.

[7] S. Mallat. A theory for multiresolution signal decomposition: The wavelet representation. *IEEE Trans. Patt. Anal. Mach. Intell.*, 11(7):674–693, 1989.

[8] T. Randen, V. Alvestad, and J. H. Husϕy. Optimal filtering for unsupervised texture feature extraction. In *Proc. Visual Communication and signal Processing*, pages 441–452, March 1996.

[9] T. Randen and J. H. Husϕy. Filtering for texture classification : a comparative study. *IEEE Trans. Patt. Anal. and Machine Intell.*, 21:291–310, April 1999.

[10] A. Teuner, O. Pichler, and B. J. Hosticka. Unsupervised texture segmentation of images using tuned matched gabor filters. *IEEE Trans. Image Processing*, 4:863–870, June 1995.