ADAPTIVE BASIS SELECTION FOR MULTI TEXTURE SEGMENTATION BY M-BAND WAVELET PACKET FRAMES

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

In this paper we propose an approach for texture feature extraction based on M-band wavelet packet frames. The features so extracted are used for segmentation of multitexture images. Standard dyadic wavelets are not suitable for the analysis of high frequency signals with relatively narrow bandwidth and also are not translation invariant. Also since most significant information of a texture often lies in the intermediate frequency bands, the present work employs an overcomplete wavelet decomposition scheme called discrete M-band wavelet packet frame (DM-bWPF), which yields improved segmentation accuracies. Wavelet packets represent a generalization of the method of multiresolution decomposition and comprise of all possible combinations of subband tree decomposition. We propose a computationally efficient search procedure to find the optimal basis based on some maximum criterion of textural measures derived from the statistical parameters of each of the subbands, to locate dominant information in each subbands (frequency channels) and decide further decomposition.

1. INTRODUCTION

The segmentation of textured images has been recognized as a difficult problem in image analysis. Of the several approaches available for texture feature extraction we focus on the signal processing approach in the present work.

Most of the texture segmentation algorithms based on signal processing techniques [1] apply the textured image to a filtering step followed by a nonlinear operation which gives an estimate of the energy. Gabor filters designed to respond to different spatial frequencies have been used in this work. However a large combination of parameters makes texture discrimination using Gabor filters computationally expensive. Randen and Husϕy [2] have examined the performance of texture segmentation schemes based on a more general class of filters including Gabor filters. Recent development of wavelet theory has provided a promising alternative and has several potential advantages over Gabor filters. The work of Chang and Kuo [3] indicates that the texture features are more prevalent in the intermediate frequency bands. Laine and Fan [4] carried out studies on texture analysis based on this concept.

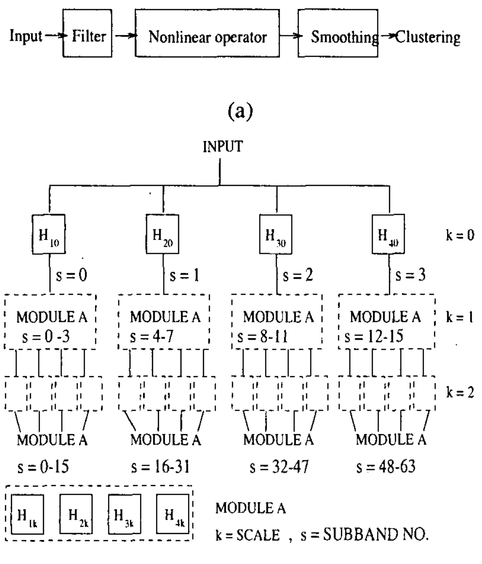
The octave band decomposition gives a logarithmic frequency resolution and are not suitable for the analysis of high frequency signals with relatively narrow bandwidth. So the main motivation of this work is to utilize the decomposition scheme based on M-band wavelets, which unlike the standard wavelet decomposition gives a mixture of logarithmic and linear frequency resolution.

Since the most significant information of a texture often appears in the middle frequency channels, and also translational invariance is desirable for accomplishing texture analysis, an M-band wavelet packet frame transform is envisaged, which corresponds to a general tree-structured filter bank, and gives an overcomplete representation. But this decomposition scheme leads to a large number of independent basis. We propose a computationally efficient and adaptive technique for finding out the optimal basis based on some maximum criterion of textural measures derived from the statistical parameters extracted from each of the subbands.

Other texture recognition technique using a non-parametric multi-scale statistical model [5] in which the joint occurrence of local features at multiple resolutions has been used to measure the similarity between textures, is reported in the literature. A common method to represent texture is to use the marginal probability densities over the outputs of a set of multi-orientation, multi-scale filters as a description of the texture, Independent Components Analysis (ICA) has been used to find out the set of filters that yield the most informative marginals [6].

Section 2 presents the analysis of the filtering technique used in the proposed work, while section 3 discusses about the extraction of features and integration of these features. Finally section 4 gives experimental results and concludes our study.

0-7803-6725-1/01/$10.00 ©2001 IEEE



(a)

(b)

Fig. 1. (a) Experimental setup. (b) Structure of DM-bWPF transform.

2. MULTISCALE WAVELET REPRESENTATION

The feature extraction scheme that we have used has a filtering stage and a subsequent nonlinear stage followed by a smoothing filter as shown in figure 1.a.

2.1. M-band wavelet filter

M-band orthonormal wavelets are a direct generalization of the 2-band orthogonal wavelets, which are able to zoom in onto narrowband high frequency components of a signal and gives better energy compaction than 2-band wavelets.

An M-band wavelet system consist of $M-1$ wavelets, $\psi_i(x)$, $i = 2, \ldots, M$ associated with the scaling function $\psi_1(x)$ [7]. The filter bank in essence is a set of bandpass filters with frequency and orientation selective properties. In the filtering stage we make use of orthogonal and linear phase M-band (M=4) wavelet following [8]. The 1-D M(=4)-band wavelet filter responses are given by ψ_i and their corresponding transfer functions are denoted by H_i, where $i = 2, \ldots, M$.

The filters $H_{ik}(\omega)$, where $i = 1, \ldots, 4$ and at level k are generated as follows,

$$H_{i,k} = H_{i,0}(2^k \omega) \quad \text{for} \quad i = 1, \ldots, 4 \tag{1}$$

Suppose $\hat{I}^k_s(\omega)$ be the Fourier transform of the input signal $I(x)$ for subband (frequency channel) s at decomposition level k. For $0 \leq s \leq 4^k - 1$ we have,

$$\hat{I}^{k+1}_{i,4s+(i-1)}(\omega) = H_{i,k}(\omega)\hat{I}^k_{i,s}(\omega) \tag{2}$$

This corresponds to a filter bank with channel filters,

$$filt^k_{i,s}(\omega) \quad |i = 1, \ldots, 4$$

from the filter bank theoretic point. The filters are given by the following recursive relation as follows,

$$\begin{aligned} filt^0_{i,0}(\omega) &= H_i(\omega), \\ filt^{k+1}_{i,4s+(i-1)}(\omega) &= H_{i,k+1}(\omega) filt^k_{i,s}(\omega) \\ &= H_{1,0}(2^{k+1}\omega) filt^k_{i,s}(\omega) \end{aligned} \tag{3}$$

For 2-D signals, the transform is obtained by the tensor product of the 1-D channel filters.

2.2. Adaptive basis selection

An appropriate way to perform the wavelet transform for texture feature extraction is to detect the most significant frequency channels and then decompose them further. This leads naturally to a tree structured wavelet transform. It is also usually redundant to decompose all the subbands in each scale to achieve the full tree of decomposition. Also for a decomposition depth of k, M^{M^k} number of basis are possible. It is quite evident that an exhaustive search to determine the optimal basis from this large set is computationally expensive and difficult procedure.

In order to avoid a full decomposition and simultaneously to find out the optimal basis, we propose an adaptive decomposition algorithm using a maximum criterion of textural measures based on the statistics extracted from each of the subbands, and identify the most significant subbands and then decide whether further decomposition of the particular channel would generate more information or not. This search is computationally efficient and enables us to zoom into any desired frequency channel for further decomposition. Energy is used as the textural measure in our work, and measures textural uniformity, meaning pixel pair repetitions.

Figure 1.b. shows a general tree structure of discrete M-band wavelet packet frame decomposition. At scale k=0, the image is first decomposed into $M \times M$ channels using all the filters H_{i0} with $i = 1, \cdots 4$. The process is repeated for each of the subbands for subsequent scales. Module A in figure 1 comprises of all the filters from H_{i0} with $i = 1, \cdots 4$. The image is first decomposed into $M \times M$ channels using the 2-D M-band wavelet transfrom, but without downsampling. Energy of each subband is then evaluated. From the M^2 subbands only those that contain appreciable energy are considered and decomposed further. In our simulation we have considered only those bands that contain more than 2% of the total energy of the

parent band. And we have further decomposed a subband if it atleast have more than 50% of the total energy of all the subbands at the current scale. This step results in a set of feature images $Feat_k(x, y)$, from which a set of feature vectors are derived.

3. COMPUTING AND INTEGRATING TEXTURE FEATURES

3.1. Local energy estimator

The objectives of the filtering and that of the local energy estimator (nonlinear operator and smoothing filter), are to transform the edges between textures into detectable discontinuities. The local energy estimator, estimates the energy of the filter output in a local region around each pixel. We have used *average absolute deviation* from the mean in small overlapping windows centering every pixel as a generalized energy definition. The local energy $Eng_k(i, j)$ around the i, j^{th} pixel is formally given as,

$$Eng_k(i,j) = \frac{1}{R}\sum_{m=1}^{W}\sum_{n=1}^{W} \mid (F_k(m,n) - \overline{F}_k(i,j)) \mid \quad (4)$$

where, W is the window size and $R = WXW$, while $\overline{F}_k(i, j)$ is the mean around the $(i, j)^{th}$ pixel and $F_k(i, j)$ is the filtered image at different scales k.

The nonlinear transform is succeeded by a Gaussian low pass (smoothing) filter $h_G(x, y)$. Formally, the feature image $Feat_k(x, y)$ corresponding to filtered image $F_k(x, y)$ is given by,

$$Feat_k(x,y) = \sum_{(a,b)\in G_{xy}} \Gamma(F_k(a,b)h_G(x-a,y-b))$$

where k correspond to different scales, $\Gamma(\cdot)$ is the nonlinear function and G_{xy} is a $G \times G$ window centered at pixel with coordinates (x, y). The size G of the smoothing or the averaging window is an important parameter. More reliable measurement of texture feature demands larger window sizes. On the other hand, more accurate localization of region boundaries requires smaller windows.

3.2. Integrating the feature images

We need to integrate these feature images to produce a segmentation. If our texture features are capable of discriminating the K texture categories (say), in a image, then the patterns belonging to each category will form a cluster in the feature space which is compact and isolated from clusters corresponding to other texture categories. We are particularly interested in the feature extraction part of this work, so we have used the simple $K - means$ clustering algorithm.

4. RESULTS

We have applied our texture segmentation algorithm to several multitexture images, in order to demonstrate the performance of our algorithm. The percentage of correctly classified pixels has been used as the segmentation quality measure in this work.

Several approaches to multichannel filtering for texture segmentation have been studied in [2]. We present a comparative performance evaluation of several approaches of texture segmentation found in the literature so far and that proposed by us with respect to the texture mosaics $Nat5b$ and $Nat5c$, of size 256 × 256 and each comprising of five different brodazt textures (figures 2a. and 2b.). Although figure 2b. looks simple, it was very difficult to segment using different techniques that are available in the literature.

Randen and Husϕy [2] compares the performance of several heuristically designed filter banks used as the feature extractors. Considering the average performance over several texture mosaics, the quadrature mirror filter (QMF) and wavelet frame approaches produce better results. Most of the heuristically designed filter banks have been reported to yield successful segmentation on several test images. But most of the them use large number of features which result in high computational complexity. Therefore optimization of the filtering operation with respect to some explicit criterion related to texture classification is desirable. Gabor filter and finite impulse response (FIR) filters were optimized following Mahalanobis and Singh (J_{MS}), Unser (J_U) and Fisher (J_F) criterion [2]. Performance of various approaches using several optimized filters, and also full rate and critically sampled wavelet filters are given in [2]. Table 1 below summarizes the results achieved using the present method and those reported in [2], for a comparative study.

methods/ filters	*test figures*	
	Nat-5b	Nat-5c
Heuristic Dyadic Gabor filter bank (d)	89.3%	77.4%
Optimized w.r.t J_U/ J_F	87.3%	73.5%
Full rate QMF $F\_2\_1\_smpl$ (d)	87.3%	79.4%
Critically sampled QMF f32d (b)	89.2%	71.9%
Proposed method M-band wavelet	97.9%	84.9%

Table 1. Performance evaluation of $Nat5b$ and $Nat5c$

In order to prove the efficacy of our algorithm we have tested it over two other 10-texture mosaics of size 256 x 512 each comprising of ten different Brodazt textures (figures 2c. and 2d). $Nat10b$ was studied by Randen and Husϕy in [2], and the results reported by them show a random performance for all approaches, so we were interested in presenting a comparative study with our result over other approaches in [2], which is given in table 2. It is quite clear from the above discussion that almost all the texture mosaics were well discriminated by our algorithm.

When comparing the performance results, complexity issues should also be taken into account. For the filtering approaches to segmentation, filtering and classification are the main contributors to the total complexity of the system. The heuristically designed dyadic Gabor filter bank approach has high filtering complexity and feature dimensionality is also large with number of features equaling 20. The optimized filtering approach has low feature count and consequently low computational complexity. But it is not applicable to an unsupervised classification problem, because for finding out the optimal parameters a first hand knowledge about the input image needs to be known. The QMF filter banks have high filtering complexities. A 40-dimensional feature extractor puts high computational load on the system [2].

One major advantage of our scheme over other methods is that even though we have made use of overcomplete wavelet representation of images, which imply large feature space (i.e. feature images of the same dimension as the input image), we have found that by intelligent selection of basis, only 7 to 11 features were sufficient for the desired segmentation of images considered in our study.

This implicates that in our method, dimensionality of the feature space can be greatly reduced, compared to the other methods reported in [2], while still maintaining a high segmentation quality.

The spatial extent of the smoothing filter that have been used in our approach ranges from 11 × 11 to 31 × 31. Simple morphological operation like median filtering can be applied to the class maps as a post-processing step to improve the segmentation results.

Another point which is mention worthy is that except for the knowledge of the number of classes present in a composite image we do not have any *a priori* knowledge about the test images, that is our scheme is completely unsupervised. While all the other approaches reported in [2] are supervised.

Four examples are by far not enough experimental data to judge which approach is best, but the tables above clearly demonstrate that the approach we have proposed is better for the images we have used.

methods/ filters	*test figures*	
	Nat-10a	Nat-10b
Heuristic Dyadic Gabor filter bank (d)	67.7%	11.6%
Optimized w.r.t J_U / J_F	64.1%	15.0%
Full rate QMF *F_2_1_smpl* (d)	58.3%	9.5%
Critically sampled QMF f32d (b)	57.7%	8.5%
Proposed method M-band wavelet	79.7%	39.2%

Table 2. Performance evaluation of $Nat10a$ and $Nat10b$

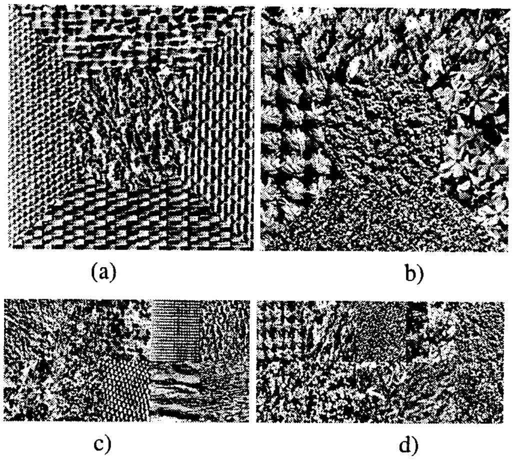

(a) b)
c) d)

Fig. 2. a) $Nat5b$ b) $Nat5c$ c) $Nat10a$ d) $Nat10b$

5. REFERENCES

[1] A. K. Jain and F. Farrokhnia, "Unsupervised texture segmentation using gabor filters," *Pattern Recognition*, vol. 24, no. 12, pp. 1167–1186, 1991.

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

[3] T. Chang and C. C. J. Kuo, "Texture analysis and classification with tree structured wavelet transform," *IEEE Transactions on Image Processing*, vol. 2, no. 4, pp. 42–44, 1993.

[4] A. Laine and J. Fan, "Frame representation for texture segmentation," *IEEE Trans. Image Process.*, vol. 5, no. 5, pp. 771–779, 1996.

[5] J. S. De Bonet and P. Viola, "Texture recognition using a non-parametric multi-scale statistical model," in *Proc. IEEE Computer Society Conference on CVPR*, Santa Barbara, California, June 1998, pp. 641–647.

[6] R. Manduchi and J. Portilla, "Independent component analysis of textures," in *Proc. 7th IEEE International Conference on Computer Vision*, Kerkyra, Greece, September 1999, pp. 1054–1060.

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

[8] O. Alkin and H. Caglar, "Design of efficient m-band coders with linear phase and perfect reconstruction properties," *IEEE Trans. Signal Processing*, vol. 43, no. 7, pp. 1579–1590, 1995.