A COMPARISON OF WAVELET TRANSFORM FEATURES FOR TEXTURE IMAGE ANNOTATION

W. Y. Ma and *B. S. Manjunath*

Department of Electrical and Computer Engineering
University of California, Santa Barbara, CA 93106

ABSTRACT

A comparison of different wavelet transform based texture features for content based search and retrieval is made. These include the conventional orthogonal and bi-orthogonal wavelet transforms, tree-structured decompositions, and the Gabor wavelet transforms. Issues discussed include image processing complexity, texture classification and discrimination, and suitability for developing indexing techniques.

1 INTRODUCTION

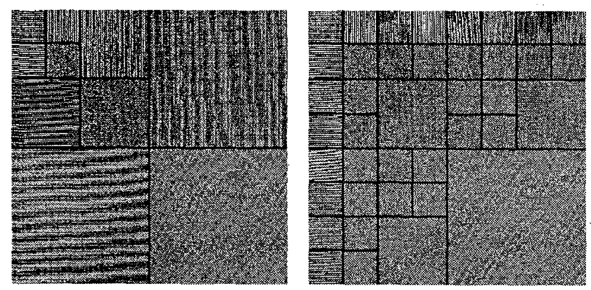
Many approaches have been suggested for texture based pattern retrieval. These range from using random field models to multiresolution techniques such as the wavelet transform. An image can be considered as the combination of different texture regions, and the image features associated with these texture regions can be used for searching and retrieving the image data. In order to make this practical, one need to address issues which involve both image feature extraction for texture analysis and efficient indexing structure design for data management [1]. Recent research on texture analysis has shown that algorithms using the multiresolution wavelet transform achieve very good performance [2], and we provide here a detailed comparison of their performance in the context of texture pattern retrieval.

We investigate the performance of different types of wavelet-transform based texture features. In particular, we consider

- orthogonal wavelet transform (OWT)
- bi-orthogonal wavelet transform (BWT)
- tree-structured decomposition using orthogonal filter bank (TOF)
- tree-structure decomposition using bi-orthogonal filter bank (TBF)
- Gabor wavelet transform (GWT)

2 WAVELET TRANSFORM FEATURES

The filter coefficients used for computing OWT and TOF are the 16-tap Daubechies wavelets [3], and for the bi-orthogonal cases (BWT and TBF) we use the 5/3 filters proposed by Le Gall [4]. Our experiments indicate that the specific set of filter coefficients is not very critical. The features obtained by OWT and BWT can be viewed as a subset of features of TOF and TBF. The difference between these two transforms is that the former decomposes only the LL band at each level whereas the latter allows decomposition in other subbands as well (LH, HL, and HH bands).



(a) (b)

Figure 1: The decomposition by (a) conventional wavelet transform. (b) tree-structured decomposition.

In [5], it is shown that a tree-structure decomposition may provide useful information to discriminate texture patterns. For some textures, the most important information for classification is often in the middle bands. Decomposition of image subbands at each level is based on energy considerations and this results in a tree structured decomposition where different patterns have different structures. However, it is observed in [5] that the energy in different bands is more stable for classification than the structure itself. For pattern retrieval applications, it is convenient to have a fixed structure. A fixed decomposition tree can be obtained by sequentially decomposing the LL, LH, and HL subbands, as shown in Figure 1. The HH band is not decomposed as this often does not lead to stable features. A fixed structure facilitates distance computations and hence database browsing.

We use three levels of decomposition of the wavelet transform. For the non-orthogonal Gabor wavelet transform, we use 6 different scales and 12 orientations/scale. Details about designing these filters can be found in [6].

The wavelet transformation involves filtering and subsampling. A compact representation needs to be derived in the transform domain for classification and retrieval. The first three moments of the energy distribution corresponding to each of the subbands at each decomposition level are used to construct the feature vector. Let the image subband

0-8186-7310-9/95 $4.00 © 1995 IEEE

be $W_{mn}(x, y)$ and mn denotes the specific subband. Then the resulting feature vector $\vec{f}$ will be $\{f_{mn}(k)\}$ where

$$f_{mn}(k) = \iint |W_{mn}(x, y)|^k dxdy, \quad k = 1, 2, 3. \tag{1}$$

This results in 12 (4×3) feature components (each of them has 3 parameters) for traditional wavelet transform, 52 ($4(1 + 3 + 9)$) for the tree-structured decomposition, and 72 (6×12) for the Gabor wavelets.

The distance measure used for comparing and retrieving the similar patterns is defined to be

$$d(i,j) = \sum_m \sum_n d_{mn}(i,j) \tag{2}$$

$$d_{mn}(i,j) = \sum_{l=1}^{3} \left| \frac{f_{mn}^{(i)}(l) - f_{mn}^{(j)}(l)}{\sigma_{mn}(l)} \right| \tag{3}$$

where i and j denote two image patterns and $\sigma_{mn}(l)$ is the standard deviation of the distribution of features $\{f_{mn}(k)\}$ in the image database.

3 PERFORMANCE EVALUATION

The image database consists of 116 different textures, each 512 x 512 pixels. A total of 1856 different images are obtained by dividing each of the 116 images into 16-128x128 sub-images.

3.1 Texture Classification

The first evaluation is based on using the entire feature vector. Given one image as the test pattern, the distance measure $d(i,j)$ is used to retrieve the most similar patterns from the database. The retrieved images are ordered according to increasing distance from the test pattern. Figure 2 shows the retrieval performance as a function of number of patterns retrieved. For example, for the Gabor wavelet features, if we retrieve the top 50 patterns, then on the average 85% (13 out of 15) of the correct textures will also be retrieved.

The graph shows that the GWT features have the best retrieval accuracy but the others are also very close. The orthogonal wavelet features (OWT and TOF) are slightly better that the bi-orthogonal ones (BWT and TBF).

3.2 Discriminating Power of the Individual Feature Components

How good are the individual feature components? An insight into this can be obtained by considering the intra-class to inter-class distance ratio DI_{mn}/DR_{mn} where

$$DI_{mn} = \frac{1}{N_1} \sum_{(\text{class of } i \neq \text{class of } j)} d_{mn}(i,j) \tag{4}$$

$$DR_{mn} = \frac{1}{N_2} \sum_{(\text{class of } i \neq \text{class of } j)} d_{mn}(i,j) \tag{5}$$

N_1 and N_2 normalize the distances. The left hand column

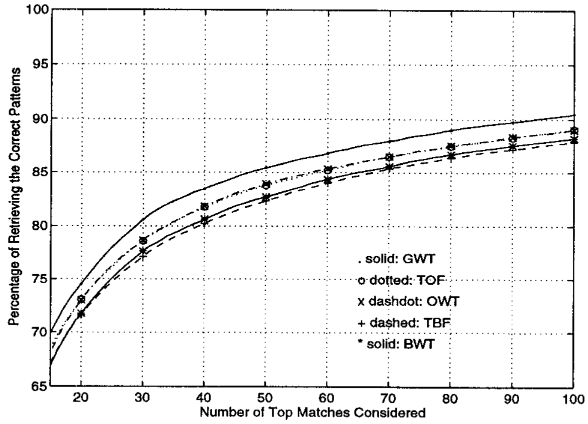


Figure 2: Retrieval performance according to number of top matches considered.

of Figure 3 shows this ratio for the different transforms, plotted as a function of the individual feature components. This information is useful in selecting more discriminating features for indexing purposes. In the figure, the feature components are ordered as follows:

OWT and BWT: features 1-4 are level 1, features 5-8 for level 2, and so on. Within each level, they are numbered in the order LL, LH, HL, and HH.

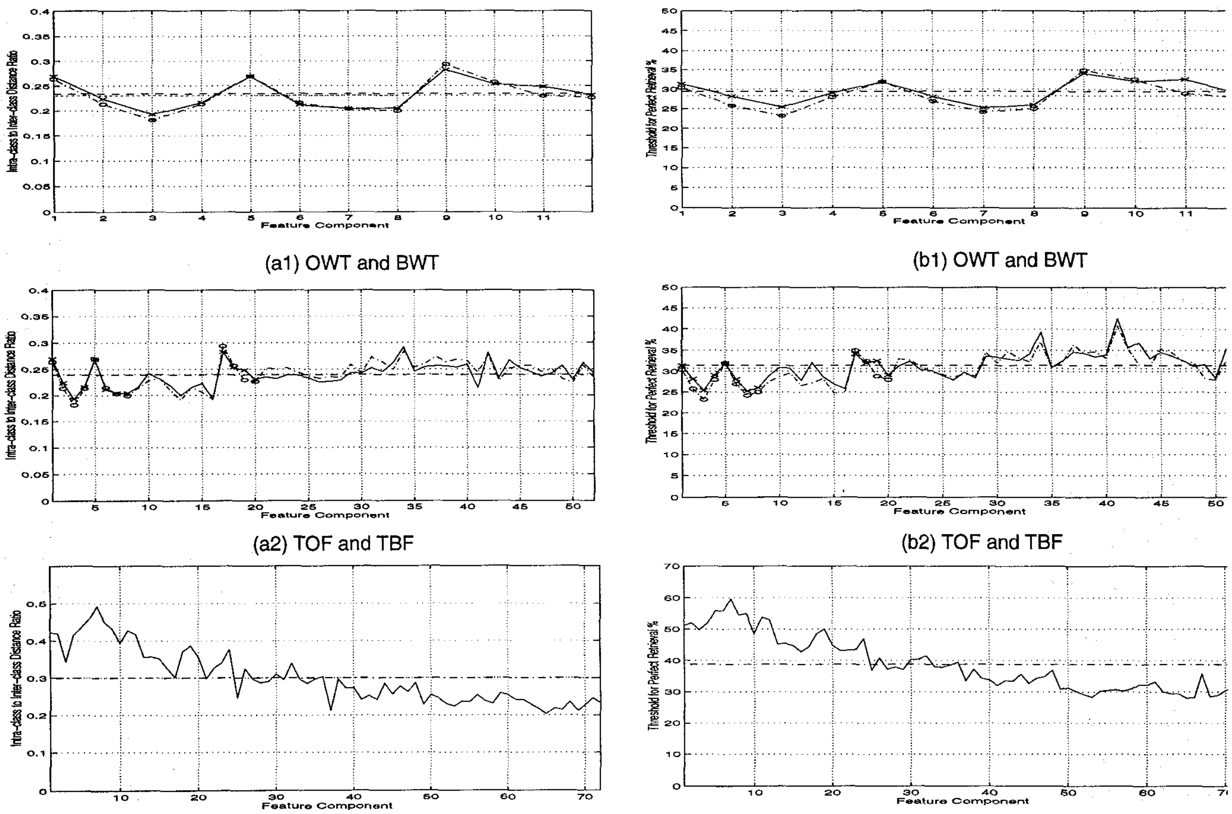
TOF and TBF: The features are ordered by decomposition level and subband. (features 1-4 for level #1 and with LL, LH, HL, and HH subbands, features 5-8 for level #2 LL band, features 9-12 for level #2 LH band, and so on).

GWT: The feature components are ordered by orientation and scale (features 1-12 for scale #1 (the lowest frequency) and increasing orientation angle, 13-24 for scale #2, and 61-72 for scale #6 (the highest frequency)).

Often in database search one is interested in finding out how much of the search space can be eliminated by using a particular feature. Suppose we want to keep all the 15 correct textures belonging to the same pattern in the set of retrieved images. The right hand column of Figure 3 illustrates this as a percentage of the total number of images in the database, for each of the feature components. For example, in the figure, it is 30% for the 50th feature for the Gabor case. This means that by keeping the top 30% of the patterns on the average ensures retaining all the correct textures.

To summarize the observations:

- In general, feature components corresponding to higher frequencies have better discriminating performance.
- At any given level of the wavelet decomposition, the LH, HL, and HH bands have better performance than the LL band. (see (a1) and (b1) in Figure 3).
- The orthogonal wavelet texture features are slightly better than the bi-orthogonal ones.



(a1) OWT and BWT

(b1) OWT and BWT

(a2) TOF and TBF

(b2) TOF and TBF

(a3) GWT

(b3) GWT

Figure 3: (a1)-(a3) show the ratio of average intra-class distance to the average inter-class distance for each feature component. (b1)-(show the threshold for retaining all 15 similar patterns from the same texture class. (a1) and (b1) are conventional wavelet transform featu (a2) and (b2) are tree-structured decomposition. The solid line is for orthogonal and dashdot line is for bi-orthogonal basis functions. (a3) ; (b3) are for the Gabor feature set.

- Even though OWT and BWT features can be considered as a subset of features in the corresponding tree-structured decomposition (TOF and TBF), they achieve almost the same performance. This is an important observation as the dimensionality of the tree-structured features is significantly larger than the conventional wavelet features.
- Decomposing the HH band in the tree-structured representation often leads to a decrease in performance, indicating that these features are not very robust.

Figure 4 shows some examples of pattern retrieval using the different texture features. In this experiment, the entire feature set is used.

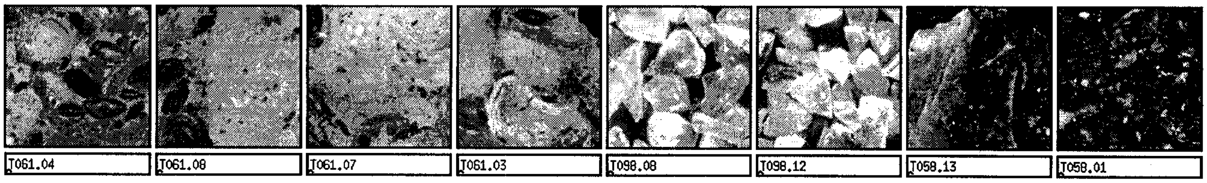
4 CONCLUSIONS

The (tree-structured) orthogonal and bi-orthogonal wavelet transforms can be efficiently performed using the FIR filter bank. The conventional orthogonal and bi-orthogonal wavelet transforms have many advantages such as lower feature dimensionality and lower image processing complexity. These transforms also facilitate efficient multiresolution storage and browsing for large image database applications. The performance of texture classification using the associated feature vector is also reasonably good and comparable to the non-orthogonal Gabor wavelet transform, which is computationally more expensive.

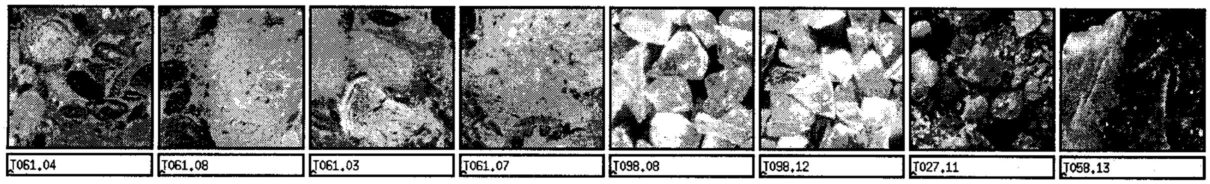
The tree-structured decomposition does not appear to significantly improve the classification and pattern retrieval performance over their conventional counterparts. At the same time, their dimensionality is much larger, making indexing more difficult.

In all our experiments, the best performance was achieved using the Gabor features. Although computationally more expensive, they are easy to interpret and offer flexibility in controlling the orientation and scale information, and are amenable for developing scale and orientation invariant features [7].

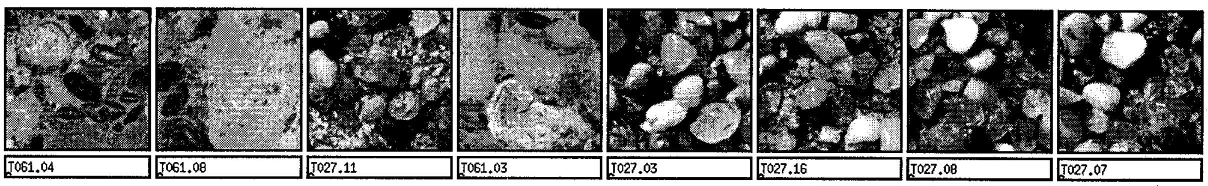
Finally, an adaptive feature selection scheme to reduce the image processing complexity for the GWT features has been suggested [8],[6]. In this scheme, a subset of Gabor filters is identified for processing a given image query. This subset is chosen based on the spectral characteristics of the



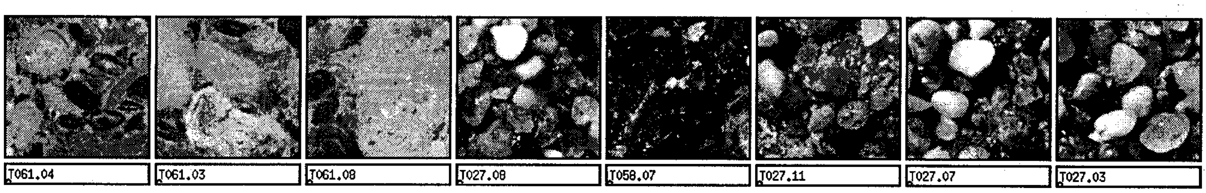
(a) Orthogonal wavelet transform



(b) Bi-orthogonal wavelet transform



(c) Tree-structured decomposition using orthogonal filter bank



(d) Tree-structured decomposition using bi-orthogonal filter bank

(e) Gabor wavelet transform

Figure 4: Retrieval examples using different wavelet transform features. The upper-left image in each block is the query pattern and the t
7 retrieved texture images are displayed from left to right with increasing distance.

image as well as the image database properties. Preliminary results are very encouraging.

Acknowledgments: We thank Dr. Michael Lightstone for providing the programs of wavelet transforms. This research was partially supported by NSF/IRI-9411330 and by NASA/NAGW-3951.

5 REFERENCES

[1] W. Y. Ma, B. S. Manjunath, "Pictorial queries: Combining feature extraction with database search," UCSB/CIPR TR 94-18.

[2] J. R. Smith, S. F. Chang, "Transform features for texture classification and discrimination in large image databases," Proc. ICIP, vol. III, pp. 407-411, Austin, TX, Oct. 1994.

[3] I. Daubechies, "The wavelet transform, time-frequency localization and signal analysis," IEEE Trans. Information Theory, vol. 36, pp. 961-1005, Sep. 1990.

[4] D. Le Gall and A. Tabatabai, "Subband coding of digital images using symmetric short kernel filters and arithmetic coding techniques," Proc. ICASSP '88, pp. 761-5, 1988.

[5] T. Chang, C.-C. Jay Kuo, "Texture analysis and classification with tree-structured wavelet transform," IEEE T-IP, vol. 2, no. 4, pp. 429-441, October 1993.

[6] W. Y. Ma, B. S. Manjunath, "Texture features for browsing and retrieval of image data," submitted for publication.

[7] G. M. Haley and B. S. Manjunath, ICIP 95 proceedings

[8] A. D. Alexandrov, W. Y. Ma, A. El Abbadi, B. S. Manjunath, "Adaptive filtering and indexing for image databases," Proc. SPIE Vol. 2420, pp. 12-23, San Jose, CA, Feb. 1995.