TEXTURE FEATURES FOR DCT-CODED IMAGE RETRIEVAL AND CLASSIFICATION*

Yu-Len Huang and Ruey-Feng Chang

Department of Computer Science and Information Engineering
National Chung Cheng University
Chiayi, Taiwan 62107, R.O.C

ABSTRACT

The multiresolution wavelet transform has been shown to be an effective technique and achieved very good performance for texture analysis. However, a large number of images are compressed by the methods based on discrete cosine transform (DCT). Hence, the image decompression of inverse DCT is needed to obtain the texture features based on the wavelet transform for the DCT-coded image. This paper proposes the use of the multiresolution reordered features for texture analysis. The proposed features are directly generated by using the DCT coefficients from the DCT-coded image. Comparisons with the subband-energy features extracted from the wavelet transform, conventional DCT using the Brodatz texture database indicate that the proposed method provides the best texture pattern retrieval accuracy and obtains much better correct classification rate. The proposed DCT based features are expected to be very useful and efficient for texture pattern retrieval and classification in large DCT-coded image databases.

The detail simulation results can be found in web page: http://www.cs.ccu.edu.tw/~hyl/mrdct/.

1. INTRODUCTION

With the growth of multimedia application and the spread of Internet, the access of digital image becomes effortless. Hence, the image content-based retrieval is essential for digital image libraries and databases. Although manual image annotations can provide a certain of help for retrieving images, it is still a complicated question to large collections of digital images. Thus, there is a new focus on computer automated approaches for image retrieval. Many approaches have been proposed for texture based image retrieval using the multiresolution techniques such as wavelet transform [1] and subband analysis [2]. Several researches have shown that algorithms using the multiresolution wavelet transform can achieve very good performance on texture analysis [3-6]. In [4], the wavelet transform can obtain much better correct classification rate and retrieval accuracy than other typical image decompositions such as DCT and spatial partitioning. Furthermore, many types of wavelet transform based texture features for image retrieval are compared in [5]. The best performance is achieved by using the Gabor wavelet transform. However, the computational effort of the Gabor wavelet transform is too expensive. The orthogonal wavelet transform can be efficiently performed using FIR filter bank. The performance of the pattern retrieval accuracy using the orthogonal wavelet transform based features is also close to that using the Gabor wavelet transform.

The discrete cosine transform (DCT) is currently the most effective and popular technique for image and video compression. It has been adopted by most emerging image coding methods including JPEG [7], H.261 [8], and MPEG [9]. All of the standards use the block-based DCT coding to achieve the higher compression ratio. That is, the DCT encoded images have been predominant in large image databases. For texture image annotation using wavelet transform features, the first step is to decompress the DCT-coded images back into spatial images and then the wavelet transform features are computed from the decompressed image data. In practice, the image decompression of inverse DCT is a time-consuming task. Consequently, an efficient extraction algorithm of DCT based texture features is necessary for diminishing the computing time of the content based retrieval system. In this paper, we propose the use of DCT features for texture analysis. Comparisons with the multiresolution wavelet transform features indicate that the proposed DCT features provide the same texture pattern retrieval accuracy without decompressing the image data.

The rest of this paper is organized as follows. Section 2 describes the extraction of the texture features based on DCT coefficients in the compressed image directly. Results of comparison with other texture features using the Brodatz texture database [10] are given in Section 3. Finally, conclusions are drawn in Section 4.

2. TEXTURE FEATURE EXTRACTION

In the DCT-based image coding, the 2-D DCT is used to map an image into a set of DCT coefficients, which are then quantized and coded. For the typical DCT based coding system, the input image is first subdivided into several pixel blocks with equal size, which are then transformed to generate the blocks of DCT coefficients. The objective of this paper is to study how to extract texture features directly from the DCT coefficients in the DCT-coded image.

2.1 Wavelet Transform Based Features and Conventional DCT Based Features

For the wavelet transform based features, decomposition of image subbands can be obtained by sequentially decomposing the LL subband. The LH, HL, and HH subbands are not further decomposed. For example, a three level decomposition results in 10 nonuniform subbands. In general, retrieving the texture pattern is measured by calculating the standard deviation and the absolute mean value of each wavelet subband. Thus a 10×2 component feature vector is produced for classification and retrieval purposes.

On the other hand, the image subbands are produced uniformly by sorting the DCT coefficients of each image block for the conventional DCT based features. For example, a total 16

*This work was supported by the National Science Council, Taiwan, Republic of China, under Grant NSC-88-2218-E-194-011.

uniform subbands will be produced from a 4 × 4 DCT coded image. As shown in Fig. 1, the DCT coefficients within the block will be stored into 16 different uniform subbands in zig-zag order. For each uniform subband, the energies are measured by calculating the variance and mean absolute value. Although the energy-based feature sets using DCT coefficients yield acceptable retrieval performance, the texture analysis using the multiresolution wavelet transform obtains much better classification and retrieval rate. However, the image decompression of inverse DCT is needed to obtain the texture features based on the wavelet transform for the DCT-coded images and videos. We hope that the features sets derived here can be generated directly in DCT domain in order to diminish the processing time. We hope that the proposed method can obtain the high retrieval performance as using the wavelet image decomposition.

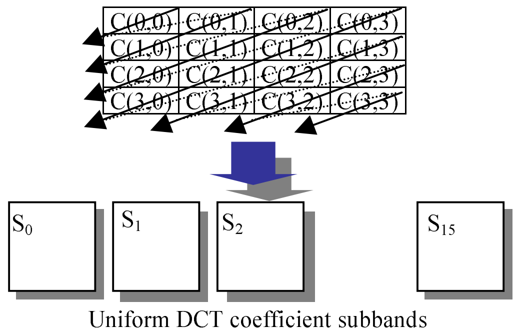

Figure 1. Sorting the coefficients of each 4×4 DCT block in zig-zag order to produce 16 uniform subbands.

2.2 Multiresolution Reordering of DCT Coefficients

It is shown that the multiresolution decomposition can provide useful information to classify texture pattern. Thus, we reorder the DCT coefficients of each image block to produce image subbands in a multiresolution decomposition-like form. First, the coefficients within each DCT image block must be denormalized. For example, the JEPG denormalization for the coefficients $T(u, v)$ within an N×N DCT block is defined as

$$C(u,v)=|T(u,v)Z(u,v)|,$$

where the $C(u, v)$ is the denormalized coefficients and **Z** is an N×N normalization array, which has been used in JPEG standard. Our experiments indicate the most important information for texture classification is often in the AC coefficients. In this paper, all of the AC coefficients are reserved to generate the feature vector in the following approach. The reconstructed mean value is obtained by the dividing the DC coefficient of each DCT block by a constant $1/N$ ($\alpha(0, 0) = 1/N$) and level shifting each divided coefficient by +128.

In order to represent the denormalized transform coefficients in a multiresolution form, we reorder the coefficients into ($3\log_2 N + 1$) multiresolution subbands for the DCT-based image with block size N×N. For a coefficient $C(u, v)$, let $2^{a-1} \le u < 2^a$ and $2^{b-1} \le v < 2^b$, where a and b are integer values. Then, the coefficient will be reordered and stored into subband $\mathbf{S}_i$, and i is computed by

$$i=\begin{cases}0 & \text{for } m=0\\ (m-1)\times 3+(a/m)\times 2+(b/m) & \text{otherwise,}\end{cases}$$

where $m = \max(a, b)$. If the coefficient $C(u, v)$ from the DCT block $\mathbf{B}(z, w)$ belongs to $\boldsymbol{S}_i$, the reordered location of $C(u, v)$ is ($z \times 2^{m-1} + u - 2^{a-1}$, $w \times 2^{m-1} + v - 2^{b-1}$). For instance, in the case of DCT block with size 8 × 8, the DCT coefficients will produce 10 multiresolution subbands. In accordance with the reordering procedure, the coefficients $C(0, 0)$, $C(0, 1)$, $C(1, 0)$, and $C(1, 1)$ of each DCT block will be stored into subbands $\mathbf{S}_0$, $\mathbf{S}_1$, $\mathbf{S}_2$, and $\mathbf{S}_3$, respectively. The coefficients $C(0, 2)$, $C(0, 3)$, $C(1, 2)$, and $C(1, 3)$ belong to $\mathbf{S}_4$ and the location can be computed by using the reordering location function. Similarly, the subbands $\mathbf{S}_5$, $\mathbf{S}_6$, $\mathbf{S}_7$, $\mathbf{S}_8$, and $\mathbf{S}_9$ can be constructed by the corresponding DCT coefficients. Fig. 2 illustrates the construction of multiresolution subbands for the coefficients within a DCT block. The decomposition by multiresolution wavelet transform and multiresolution reordered DCT are shown in Fig. 3. It is clearly that the multiresolution reordered DCT subbands contain the structure similar to that of the wavelet subbands.

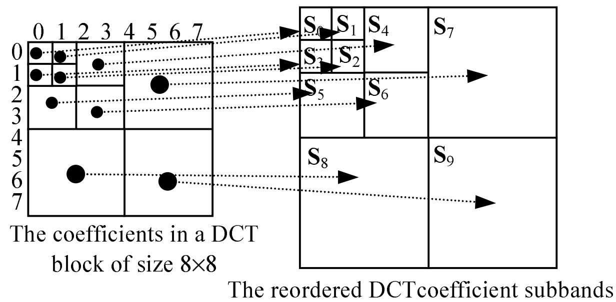

Figure 2. A reordering example for an 8×8 DCT block.

2.3 Texture Feature Representation

In the multiresolution wavelet transform decomposition, the mean and the standard deviation corresponding to each of the DCT coefficient subbands are used to construct the feature vector. Let the coefficient subband be $S_i(x, y)$ and i denotes the specific subband. The mean value μ_i and the standard deviation σ_i are computed as

$$\mu_i=\iint|S_i(x,y)|\,dxdy \quad \text{and} \quad \sigma_i=\sqrt{\iint\left(|S_i(x,y)|-\mu_i\right)^2 dxdy}\,.$$

The texture feature vector will be constructed using μ_i and σ_i as feature components. The dimension of the feature vector for the DCT-based images of block size N×N is $K = 2 \times (3\log_2 N + 1)$. That is, the feature vector $\bar{f}$ is formed as

$$\bar{f}=\left[\mu_0,\sigma_0,\mu_1,\sigma_1,\mu_2,\sigma_2,\ldots,\mu_{\frac{K}{2}-1},\sigma_{\frac{K}{2}-1}\right].$$

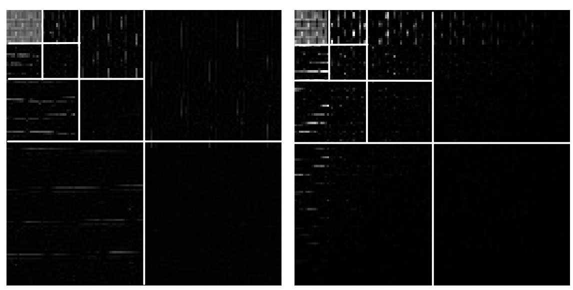

Figure 3. The decomposition by (a) multiresolution wavelet transform and (b) multiresolution reordered DCT.

2.4 Texture Classification and Pattern Retrieval Feature Distance Measure

Given two image texture patterns p_1 and p_2, the distance measure used for comparing and retrieving the patterns is defined as

$$d(p_1, p_2) = \sum_{k=0}^{K-1} \left| \frac{f_k^{p_1} - f_k^{p2}}{\alpha(f_k)} \right|,$$

where $\vec{f}^{p_1}$ and $\vec{f}^{p_2}$ represent the the feature vector of p_1 and p_2, respectively. The term $\alpha(f_k)$ is the standard deviation of the respective features in the image database. That is used to normalize the feature components. The diagram of the texture pattern retrieval scheme using the proposed multiresolution reordered DCT features is shown as Fig. 4.

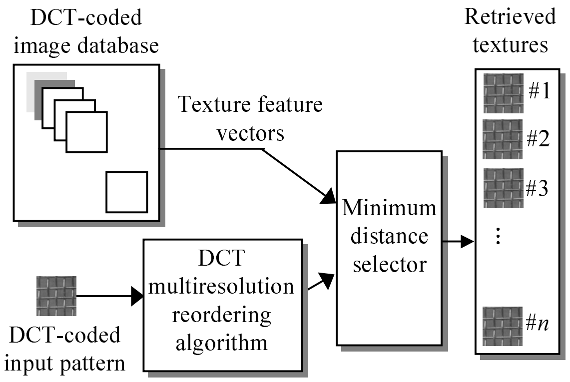

Figure 4. Texture pattern retrieval using the proposed multiresolution reordered DCT features.

3. SIMULATION RESULTS

The image database used in this paper consists of 112 different 512 × 512 texture images from the Brodatz Album [10]. For the orthogonal wavelet transform, we use the 16-tap Daubechies wavelets [1] as the filter coefficient to complete the 3 level wavelet decomposition and then produce 10 wavelet subbands. In the simulations, The size of DCT image blocks is 8×8. Hence, the proposed multiresolution reordered DCT will generate 10 DCT coefficient subbands.

3.1 Texture Classification Results

Each of the texture image is divided into 20 randomly positioned textures. Finally, there are produce 2240 total subimages. For the purpose of comparisons, we use 10 training sets with different size. For the training set i (i = 1, 2, 3,, 10), i texture cuts were selected from each class. The remaining 10 texture cuts from each class were used as the test set. For the orthogonal wavelet transform subband (OWT) and the multiresolution reordered DCT subband (MRDCT) decompositions, the 20 term feature vectors for each texture class are generated by averaging the feature vectors that belong to the same texture class. We compare the classification performance of 3 different dimensions (8, 14, and 20) of feature components for the OWT and MRDCT. Both the OWT-8 and MRDCT-8 features use the lowest 4 frequency subbands ($\mathbf{S}_0$, $\mathbf{S}_1$, $\mathbf{S}_2$, and $\mathbf{S}_3$) to generate the feature vectors with length 8 (4×2). Likewise, the OWT-14 and MRDCT-14 generated the 14 (7×2) term feature vectors by using the subbands $\mathbf{S}_0$, $\mathbf{S}_1$, $\mathbf{S}_2$, $\mathbf{S}_3$, $\mathbf{S}_4$, $\mathbf{S}_5$, and $\mathbf{S}_6$. The OWT-20 and MRDCT-20 adopt all the subbands to generate feature vectors. The classification performance of the energy-based feature sets is shown in Fig. 5. In these experiments, the better classification performance was reached using the proposed DCT feature sets.

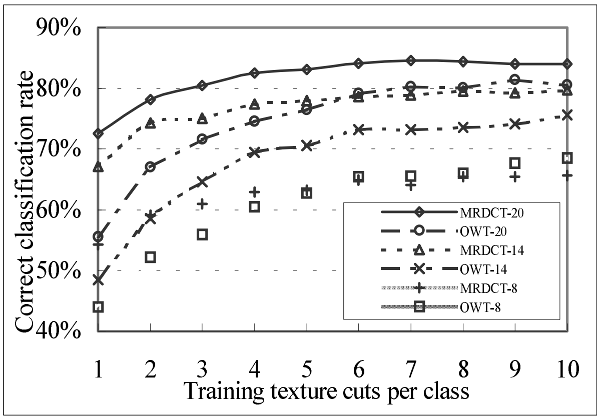

Figure 5. Classification performance according to training texture cuts per class

3.2 Pattern Retrieval Results

Each of the texture images is further divided into 16 non-overlapping 128×128 subimages for pattern retrieval simulations. Thus 1792 different texture images are generated. The test texture patterns are randomly selected form the texture images. We compare the three image decompositions that are used to produce the feature vectors: OWT, the conventional DCT (CDCT), and the proposed MRDCT. Notice that the coefficients of the conventional DCT are sorted in the zig-zag order to produce 64 subbands to generated feature vectors.

The feature vectors of the test query patterns p_q are produced using the three methods. The distance measure $d(p_q, p_i)$ is computed for each texture image in the database, where p_i is a query texture from the database. The retrieved texture images are ordered according to increasing distance from the test pattern. The ideal situation of retrieving is all the top 15 matches are from the same large texture image. The performance is measured in terms of the average correct retrieval rate which is defined as the average percentage of pattern belong to the same Brodatz image (not includes the query pattern) as the query pattern number of top matches considered. Note that the query pattern is always top one because the distance is 0.

Fig. 6 shows the retrieval performance as a function of number of patterns retrieved at the same feature vector length 20 (10×2) for the various texture features. In the conventional DCT features, only the first 10 subbands from the 64 conventional DCT subbands are used to produce a 10×2 component feature vector (denotes CDCT-20) in this comparison. It can be observed that the use of the proposed features achieves the better retrieval accuracy than the orthogonal wavelet transform in most cases. Moreover, we compare the retrieval accuracy using the MRDCT features and the CDCT feature vector with different lengths 20, 32, 64, and 128. In the Fig. 7, we can find that the MRDCT features also provide the best retrieval performance than the CDCT with the higher dimension feature vector. The retrieval results using the OWT and the proposed method at the three different dimension of feature vector are shown in Fig. 8. From the simulation results, we find that the proposed method has very good performance for texture pattern retrieval.

The detailed simulation results and retrieval examples can be found in web page: http://www.cs.ccu.edu.tw/~hyl/mrdct/.

4. CONCLUSIONS

In this paper, the texture-based classification and image retrieval are performed using the energy-based features from the wavelet transform and DCT decompositions. The multiresolution wavelet transforms have been shown that they are able to achieve high correct retrieval rate at the lower feature dimensionality. However, the texture features will be produced after the DCT decompression and the wavelet decomposition for DCT-based compressed images. Thus in this paper, we proposed a texture based analysis which can extract texture features directly from the DCT coefficients. We showed that the classification and retrieval performances of the proposed method and the method using the conventional wavelet transform are very close at the same feature dimension. We also showed that the MRDCT features achieve the best retrieval performance in comparing with the conventional DCT method using several larger feature dimensions. From the experimental results, it shows that the proposed MRDCT features are expected to be very useful and efficient for texture pattern retrieval in large DCT coded image databases. In future work, we will extend the proposed method to real image and video database applications.

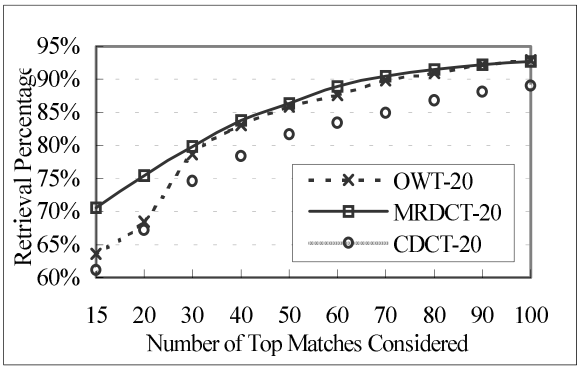

Figure 6. Retrieval performance according to the number of top matches considered (length of feature vector is 20)

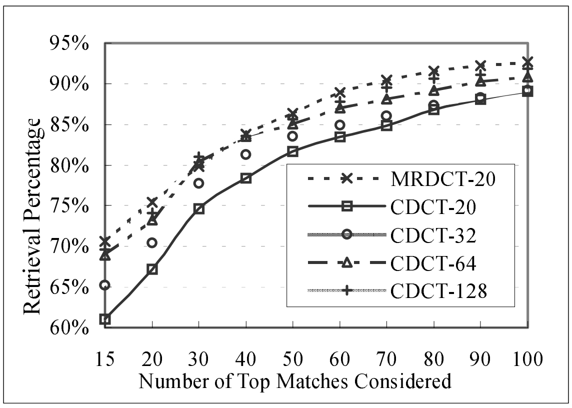

Figure 7. Retrieval performance according to the number of top matches considered (CDCT feature vectors with different lengths)

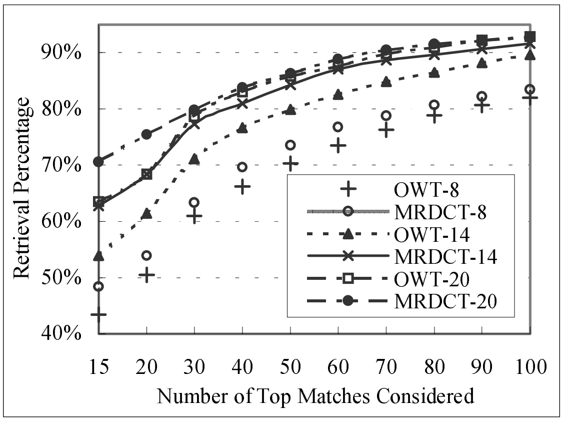

Figure 8. Retrieval performance according to the number of top matches considered (OWT and MRDCT feature vectors with variable length).

5. REFERENCES

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

[2] D. Le Gall and A. Tabatabai, "Subband coding of digital images using symmetric short kernel filters and arithmetic coding techniques," in *Proc. ICASSP*, Apr. 1988, pp. 761-764.

[3] T. Chang and C. J. Kao, "Texture analysis and classification with tree-structured wavelet transform," *IEEE Trans. Image Processing*, vol. 2, no. 4, pp. 429-441, Oct. 1993.

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

[5] W. Y. Ma and B. S. Manjunath, "A comparison of wavelet transform features for texture image annotation," in *Proc. ICIP*, vol. II, Oct. 1995, pp. 256-259.

[6] B. S. Manjunath and W. Y. Ma, "Texture feature for browsing and retrieval of image data," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 18, no. 8, pp. 837-842, Aug. 1996.

[7] G. K. Wallace, "The JPEG still picture compression standard," *Commun. of the ACM*, vol. 34, no. 4, pp.31-44, Apr. 1991.

[8] M. Liou, "Overview of the px64 kbits/s video coding standard," *Commun. of the ACM*, vol. 34, no. 4, pp.59-63, Apr. 1991.

[9] D. L. Gall, "MPEG: A video compression for multimedia applications," *Commun. of the ACM*, vol. 34, no. 4, pp.47-58, Apr. 1991.

[10] P. Bordatz, *Texture: A Photographic Album for Artists and Designers*. New York: Dover, 1966.