

IMAGE RETRIEVAL ON UNCOMPRESSED AND COMPRESSED DOMAINS*

Ruey-Feng Chang, Wen-Jia Kuo and Hung-Chi Tsai

Department of Computer Science and Information Engineering
National Chung Cheng University
Chiayi, Taiwan 621, R.O.C.
E-mail: rfchang@cs.ccu.edu.tw

ABSTRACT

With the rapid growth of Internet and multimedia system, digital images have been enormously used in many applications now. The way of how to retrieve the image we want from a large image database correctly plays a more important role. Currently, most proposed methods for image retrieval can only be applied with images in the same representing domain. In this paper, we propose a novel methodology that extracts the feature image directly from its original domain without any decompressing procedures. The feature is further modified to achieve the rotation-invariant and scaling property. Our proposed method can be used for image retrieval on the three most frequently adopted types of images nowadays: spatial, DCT compressed, and wavelet-compressed images. Experimental results show that our proposed method is quite effective not only for the performance but also is not afraid of scale and rotation.

1. INTRODUCTION

In the past few years, the general problems of content-based image retrieval have been widely studied [1]-[9] in many researches such as IBM's QBIC (Query By Image Content) system [2], QVE (Query Visual Example) system [5], *etc.* Most of the systems extract image features including color, texture, shape, and sketch to retrieve images in the spatial domain. However, these systems are not practical nowadays because many image compression methods are developed in different domains including JPEG [10]-[11] and JPEG-2000 [12]. If one still wants to use these spatial-based image retrieval systems to retrieve images, these compressed images should be first decompressed into uncompressed images. Shneier and Abdel-Mottaleb [6] proposed a method that can retrieve images directly on JPEG images without decompressing JPEG images. However, the system can not apply their method directly on uncompressed images because their source images are limited in compressed images. Meanwhile, the system can not apply their method on different kind of compressed images such as JPEG-2000 images.

In this paper, we propose a method that can retrieve images from different kinds of domains. We use a unified approach on raw image data, JPEG image, and wavelet-compressed image. Feature images are extracted from their original domains without decompressing the image no matter whether the image is compressed by DCT or wavelet transform. Scale-invariant property is then achieved by normalizing feature images to a predefined size. Then, the wavelet transform is performed for these normalized feature images and the energy of each high subband is used as the feature. Meanwhile, the proposed feature is further modified into a rotation-invariant feature. In the experiments, our image database contain over 1000 images which consist of raw image data, JPEG image, and wavelet-compressed image. The results show that our method can retrieve images correctly.

The organization of this paper is as follows. In section 2, we describe the proposed method for image retrieval. The experiments and results are shown in section 3. Finally, conclusions are given in section 4.

2. IMAGE RETRIEVAL IN VARIOUS DOMAINS

In order to achieve the purpose of directly retrieving the various types of images in their original domains, the images are preprocessed in their own domains to extract feature images. For the same image, the feature images obtained in different domains will be the same or similar. The preprocessing of various images is shown in Fig. 1.

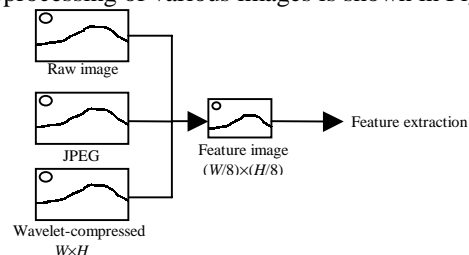


Figure 1: The feature image is directly obtained from various domains.

2.1 Feature Image Extraction for Various Image Domains

In this section, we briefly introduce the methods of extracting feature images directly from various domains. The main purpose is to find a suitable feature image

* This work was supported by the National Science Council, Taiwan, Republic of China, under Grant NSC-89-2213-E-194-025.

extracted from various domains of the same image with the same or similar properties.

2.1.1 Raw Images

In order to achieve the same feature image as the JPEG image, raw image is divided into 8×8 blocks and the mean of each block is then assigned to be the pixel value of the feature image. Thus, the pixel value of the feature image is

$$FI_{x,y} = \frac{1}{64} \sum_{j=0}^7 \sum_{i=0}^7 OI_{8x+i,8y+j}, \quad (1)$$

where $FI_{x,y}$ is the pixel value of feature image with coordinate (x, y) , and $OI_{x,y}$ is the pixel value of original image with coordinate (x, y) . We can find that the size of feature image is $1/64$ of original image size.

2.1.2 JPEG Images

Because the feature image of a raw image is composed of the mean value of each 8×8 block, the mean value of each block in the JPEG image is then directly extracted from its DC coefficient as the feature. The result can be easily inferred as follows:

$$\begin{aligned} & F(0,0) \\ &= \frac{c(0)c(0)}{4} \sum_{x=0}^7 \sum_{y=0}^7 f(x,y) \cos\left(\frac{(2x+1) \times 0 \times \pi}{16}\right) \cos\left(\frac{(2y+1) \times 0 \times \pi}{16}\right) \\ &= \frac{1}{16} \sum_{x=0}^7 \sum_{y=0}^7 f(x,y) \\ &= 4 \times \frac{1}{64} \sum_{x=0}^7 \sum_{y=0}^7 f(x,y) \\ &= 4 \times M, \end{aligned} \quad (2)$$

where $F(0,0)$ and M are the DC coefficient and mean value of the corresponding block. For the reason that a level shifting by -128 gray levels in the JPEG encoding, the real mean value of the block is $\left[\frac{1}{4}F(0,0)+128\right]$. The real mean

values of all blocks are assigned to be the pixel values of the feature image. The size of the feature image is still $1/64$ of original JPEG image size because the DCT block size is 8×8 .

2.1.3 Wavelet-compressed Images

For a wavelet-compressed image, feature image is extracted from the low-low band of the wavelet-compressed. If the one-level wavelet decomposition is used in the wavelet-compressed image, the low-low band subimage will approximate to the scaled original image. Thus, the mean value of each 4×4 block in the low-low band subimage is assigned to be the pixel value of the feature image. The pixel value of the feature image is

$$FI_{x,y} = \frac{1}{16} \sum_{j=0}^3 \sum_{i=0}^3 LL_{4x+i,4y+j}, \quad (3)$$

where $FI_{x,y}$ is the pixel value of feature image with coordinate (x, y) , and $LL_{x,y}$ is the pixel value of low-low

band image with coordinate (x, y) . The size of feature image here is $1/64$ of the original wavelet-compressed image size. If the wavelet-compressed image is compressed by three-level wavelet decomposition, then the image should be reconstructed back to the one-level wavelet decomposition first.

The feature images will be the same if they are extracted from the raw image and the JPEG image of the same image. Moreover, the mean squared error (MSE) between feature images generated from the raw image and from the wavelet-compressed image is quite small. For example, the MSE of feature images extracted from the raw image and one-level wavelet-compressed image for the image Lenna is 1.50. Note that the feature images will be normalized to a predefined size. This normalized step ensures the scale-invariant property of the proposed method.

2.2 Feature Extraction

In this section, features are then extracted from the feature images obtained from different domains for image retrieval. First, the feature image is decomposed by three-level wavelet decomposition to 10 nonuniform subbands. In general, standard deviation, absolute mean, and the energy of each wavelet subband can be used as the features to distinguish different images. In the work, we use the energy of each wavelet subband, except the LL_3 subband, as the texture features. Let the wavelet coefficients of subband i be $S_i(x, y)$. The energy of the wavelet subband i is defined as

$$E_i = \sum_{x=0}^{W-1} \sum_{y=0}^{H-1} (S_i(x, y))^2, \quad (4)$$

where W and H are the width and height of the subband i , respectively. The feature vector FV_M of an image M is

$$FV_M = (E_{LH_1}, E_{HL_1}, E_{HH_1}, E_{LH_2}, E_{HL_2}, E_{HH_2}, E_{LH_3}, E_{HL_3}, E_{HH_3}). \quad (5)$$

In this work, the distance measure d_{ij} between images i and j is defined as

$$d_{ij} = \sum_{l=1}^3 w_l \left((E_{LH_l}^i - E_{LH_l}^j)^2 + (E_{HL_l}^i - E_{HL_l}^j)^2 + (E_{HH_l}^i - E_{HH_l}^j)^2 \right), \quad (6)$$

where w_l is the weighting coefficient for level l . Because the subband image size in the higher level is much smaller than that in the lower level, the total energy value in the higher level is much smaller than that in the lower level. A larger weighting value is used for the higher levels to normalize the total energy values in different levels. In this work, the weighting coefficients in three levels are set as $(w_1, w_2, w_3) = (1, 2, 4)$, which are inversely proportional to their subband image sizes. The retrieving result is to select images from the database with smaller distance between the query image and image in the database.

In the next section, the feature will be modified to satisfy both rotation-invariant and scale-invariant properties.

2.3 Rotation-Invariant Feature

There is a straightforward and simple method to make a feature without rotation-invariant properties become a new rotation-invariant feature. First, the image is rotated by various possible angles, from 0° to 360° , to generate several rotated images. The feature of each rotated image is extracted and then all the features are summed up. Let the feature of rotated image with i degree be $F(RI_i)$. Then, the rotation-invariant feature F_{360} can be defined as

$$\sum_{i=0}^{K-1} F(RI_{360/i}),$$

where K is the total number of rotated images. However, it is time consumption to rotate the images for each degree. We must reduce the total number of rotated images. In fact, we find that the rotated images with 0° to 90° are sufficient if the feature has the symmetric property. That is, we will use the rotation-invariant feature $F_{90} = \sum_{i=0}^{K-1} F(RI_{90/i})$. The execution time of

the feature F_{90} is only the 1/4 of the time for the feature F_{360} . From Fig. 2 and Table 1, we can find that the feature has symmetric property at 90° and 180° . The total energy in the rotated angle range between 0° and 180° is equal to that between 180° and 360° . Moreover, the total energy in the rotated angle range between 0° and 90° is equal to that between 90° and 180° .

In this paper, the Shen and Sethi's rotation method [13] is adopted to rotate the feature image. The method can perform arbitrary angle rotation in the uncompressed domain. In this three-pass rotation algorithm, the rotated image RI with θ degrees for an image I is decomposed into three steps:

$$RI = I \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} = I \begin{bmatrix} 1 & 0 \\ -\tan\frac{\theta}{2} & 1 \end{bmatrix} \begin{bmatrix} 1 & \sin\theta \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -\tan\frac{\theta}{2} & 1 \end{bmatrix}. \quad (7)$$

This method skews the image along the horizontal direction in the first step. At the second step, the skewed image is then skewed along the vertical direction. Finally, another skew in the horizontal direction is performed to finish the rotation.

Table 1: Total Energy in different subbands and degree ranges.

Subband Degree range	LH	HL	HH
0~360	1695329	1700144	672135
0~180	835034	850746	329291
0~90	439178	432370	168374

After rotating the image, the rotated image should be clipped into a rectangle image. To avoid losing too much information of the feature image, we need to find a suitable way to achieve the maximal size of the subimage. As shown in Fig. 3, the height and width of the subimage are both equal to $\min(H, W)/\sqrt{2}$, where H and W are the height and width of the feature image. Finally, directly

sum these features generated from these rectangle subimages to obtain a new rotation-invariant feature.

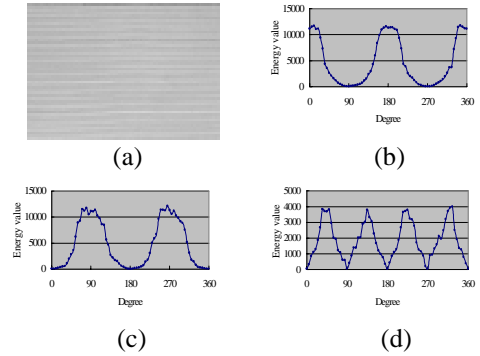


Figure 2: (a) The tested 640×480 texture image. (b) LH subband energy values. (c) HL subband energy values. (d) HH subband energy values.

3. EXPERIMENTS AND RESULTS

In this section, we show some experimental results and evaluate the performance of our proposed techniques. In our experiments, we use the rotation-invariant features to retrieve images from the image database. In our image database, there are over 1000 images consists of raw images, JPEG images, and wavelet-compressed images. The size of images in the database is 640×480 . Some experiments are performed in our retrieval system to show the properties of our proposed method. In order to evaluate the performance of our proposed retrieval system, we define the correct ratio C_r as

$$C_r = \frac{N_c}{N_q} \times 100\%, \quad (8)$$

where N_c is the number of correct retrieval and N_q is the number of query images. The retrieval system is better if the C_r value is larger.

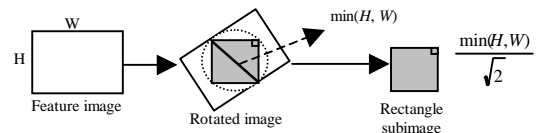


Figure 3: A rectangle subimage is extracted from the rotated feature image.

3.1 Experimental Results for Querying by the Existed Image in Database

In this section, all the query images are already existed in the image database. The amount of query images is 20. The retrieval is considered as an accurate retrieval only when the desired image in the database is ranked at the first. The C_r value for this kind of retrieval is 100%.

3.2 Experimental Results for Querying by Images in Various Domains

In this experiment, the query images are inserted into the database at various formats including raw image data, JPEG image, and wavelet-compressed image. The JPEG

format images are used as the query images. The retrieval is considered correct if all formats of the query image are ranked at top five. The C_r is 100% for the 20 experiment query images. We can find that the proposed retrieval system is effective even the queries are performed by images with various domains.

3.3 Experimental Results for Querying Images with Various Angles

In this experiment, we will show the rotation-invariant property of the proposed method. We use the non-rotated images as the query images. These query images will be rotated by the following angles: 5° , 10° , 45° , 90° , and 200° . These rotated images are then inserted into the image database. We define that the SCORE is the number of rotated images, which are retrieved at top ten. The retrieval is correct if the SCORE is larger or equal to 2. The C_r value here is 95%.

3.4 Experimental Results for Querying Images with Various Sizes

In this experiment, the query images will be scaled by the following factors: $\times 2$ and $\times 3$. These scaled images are inserted into the image database. We define that the retrieval is correct if all the scaled images can be retrieved at top ten. The correct ratio value of C_r value in this case is 80%.

3.5 Experimental Results for Querying by Images outside the Database

Here, the query images are not inserted into the image database. Fig. 4 shows an example for this experiment.

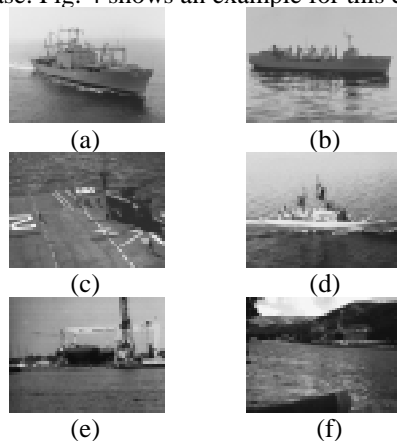


Figure 4: The result of querying the image not in the image database. (a) the query image. (b)-(f) the top five similar images.

4. CONCLUSIONS

In this paper, we proposed an image retrieval system using the wavelet coefficients as the index to retrieve images in various domains. Feature images are extracted directly in their original domains. Then, three-level wavelet decomposition is used for the feature images to generate

the features for image retrieval. Symmetric property of the features is further used to reduce the computation time when extract features from the feature images. In the experiments, we evaluated the proposed method in the image database, which contains over 1000 images with raw images, JPEG images, and wavelet-compressed images. Experimental results show that our method can retrieve images correctly with scale-invariant and rotation-invariant properties.

For future works, we can try to find some features to increase the correct ratio of retrievals. Meanwhile, the computation time for feature extraction can be further reduced to make the retrieval system more flexible.

5. REFERENCES

- [1] V. Gudivada and V. Raghavan, "Content-based image retrieval systems," *IEEE Computer*, vol. 28, no. 9, pp. 18-22, Sept. 1995.
- [2] M. Fickner, H. Sawhney, W. Niblack, J. Ashkey, Q. Huang, B. Dom, M. Gorkani, J. Hafner, D. Lee, D. Petkovic, D. Steele, and P. Yanker, "Query by image and video content: The QBIC system," *IEEE Computer*, vol. 28, no. 9, pp. 23-32, Sept. 1995.
- [3] H. J. Zhang, C. Y. Low, S. W. Smoliar, and J. H. Wu, "Video parsing, retrieval and browsing: An integrated and content-based solution," in *Proc. ACM Multimedia*, 1995, pp. 15-24.
- [4] D. H. Huang and C.L. Huang, "A content-based image retrieval system," in *IPPR Conference on Computer Vision, Graphics, and Image Processing*, 1996, pp. 259-266.
- [5] T. Tako, T. Kurita, N. Otsu, and K. Hirata, "A Sketch retrieval method for full color image database: Query by visual example," in *Proc. 11th International Conference on Pattern Recognition*, 1992, pp. 530-533.
- [6] M. Shneier and M. Abdel-Mottaleb, "Exploiting the JPEG compression scheme for image retrieval," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 18, no. 8, pp. 849-853, Aug. 1996.
- [7] Tat-Seng Chua, Swee-Kiew Lim and Hung-Keng Pung, "Content-based Retrieval of Segmented Images," in *Proc. ACM Multimedia*, 1994, pp. 211-218.
- [8] Rajiv Mehrotra, James E. Gray, "Feature-based retrieval of similar shapes," in *9th International Conference on Data Engineering*, 1993, pp. 108-115.
- [9] Hsin-Chih Lin, Ling-Ling Wang, and Shi-Nine Yang, "Color image retrieval based on hidden Markov models," *IEEE Trans. Image Processing*, vol. 6, no. 2, Feb. 1997, pp. 332-339.
- [10] G. K. Wallace, "The JPEG still picture compression standard," *Commun. of the ACM*, vol. 34, no. 4, pp. 30-44, Apr. 1991.
- [11] W. Pennebaker and J. Mitchell, *JPEG still image data compression standard*, Van Nostrand Reinhold, New York, 1993.
- [12] ISO/IEC JTC 1/SC 29/WG 1 N 751, "Coding of Still Pictures," Mar. 1998.
- [13] B. Shen and I. K. Sethi, "Block-based manipulations on transform-compressed images and videos," *Multimedia Systems*, pp. 113-124, 1998.