

The image retrieval based on transform domain

Wei-bin Fu

Hainan University
College of Information Science and Technology
Haikou Hainan, China
64043003@qq.com

Meng-xing Huang

Hainan University
College of Information Science and Technology
Haikou Hainan, China
Huangmx09@163.com

Jing-bing Li*

Hainan University,
College of Information Science and Technology
Haikou Hainan, China
Jingbingli2008@hotmail.com

Yi-Cheng LI

Hainan University,
College of Information Science and Technology
Haikou Hainan, China
57849036@qq.com

Abstract—Due to the most of popular algorithms based on image retrieval have a large calculation. While the method of DCT domain have less calculation, and compatible with international popular standard of data compression (JPEG、MPEG、H261/263). So this paper proposes an image retrieval method based on DCT compression. First, we can extract the feature vector of the image through the DCT transform, and then set up characteristic databases, finally through matching the feature vector of the normalized correlation coefficient (NC) to realize the image retrieval. The results show that this method can not only reduce the database storage space occupied, but also can effectively resist the conventional attacks and geometrical attacks, has strong robustness.

Keywords—image retrieval; DCT; feature vector; normalized correlation coefficient; robustness.

I. INTRODUCTION

With the rapid development of Internet, multimedia technology, and computer technology, people began to use a variety of advanced technology to gather and produce various types of multimedia data, including text, images, sound, video, etc. Image as a rich content multimedia data, has become an irreplaceable important network information resources, it contains more information than the text. In the face of huge amounts of image data, how to better implement quickly and accurately to retrieve the information that users need, has become the urgent problem to solve [1][2].

In the 80 s, although the multimedia technology is developing rapidly, and the image acquisition, creation, compression, storage technology has made remarkable achievements, but the management of the image information has not been given enough attention. In the early 90 s, with the emergence of large-scale digital image library makes the image data is growing fast. And the image data is transmitted through the network to all over the world. Therefore, how rapidly and efficiently from the vast amounts of image data to retrieve the

(4) At present most popular algorithms, they all have the shortcomings of great amount of calculation.

information you need is a current important problem in many applications [3].

The traditional information retrieval is based on the numerical/character. It does not objectively reflect the diversity of image content. Its mathematical model, system structure, such as inquiry mode and the user interface also does not have effective management and the ability to retrieve the image data [4]. In order to realize automatic and intelligent Image query and management mode, and achieve a single intervention management work. A new image retrieval technology, content-based Image Retrieval technology (CBIR, Content - -based Image Retrieval) was proposed and developed rapidly. CBIR based on computer vision and image understanding theory, combining the artificial intelligence, object-oriented technology, cognitive psychology, database and other multi-disciplinary knowledge. Image content description is no longer rely on manual annotation, but with the help of visual features of automatically extracted from the image, the retrieval process is no longer a keyword matching, but the similarity matching between the visual features[5][6][7].

However, the current problems still existing in content-based image retrieval are as follow:

(1)The image retrieval Based on image color feature index of the characteristics of the main problems is a person of color visual perception and consideration is still not enough.

(2)The image retrieval Based on image texture feature index of the main problems is that a variety of methods to choose texture feature set depends on the specific texture image.

(3) While the image retrieval based on shape feature, automatic extraction of shape boundary has been the main problem in image processing field for many years. In the present retrieval systems mostly adopt the way of manual outline. To extract shape feature is a very heavy work, and mass image data for this problem will appear more prominent.

In view of the above problems, this paper proposes an image retrieval method based on DCT compression.

First, we can extract the feature vector of the image through the DCT transform, and then set up characteristic databases, finally through matching the feature vector of the normalized correlation coefficient (NC) to realize the image retrieval. The algorithm is simple, and can greatly reduce the database storage space, has a certain ability to resist conventional and geometric attacks[8][9][10].

II. BASIC THEORY

A. Two-dimensional discrete cosine and reverse transformation formula

A $M \times N$ matrix A two-dimensional discrete cosine transform (DCT) is formula is as follows:

$$F(u, v) = c(u)c(v) \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) \cos \frac{\pi(2x+1)u}{2M} \cos \frac{\pi(2y+1)v}{2N}$$

$$u = 0, 1, \dots, M-1; v = 0, 1, \dots, N-1 \quad (1)$$

$$c(u) = \begin{cases} \sqrt{1/M} & u = 0 \\ \sqrt{2/M} & u = 1, 2, \dots, M-1 \end{cases} \quad (2)$$

$$c(v) = \begin{cases} \sqrt{1/N} & v = 0 \\ \sqrt{2/N} & v = 1, 2, \dots, N-1 \end{cases} \quad (3)$$

Two-dimensional discrete cosine transformation (IDCT) formula is as follows:

$$f(x, y) = \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} c(u)c(v)F(u, v) \cos \frac{\pi(2x+1)u}{2M} \cos \frac{\pi(2y+1)v}{2N}$$

$$x = 0, 1, \dots, M-1; y = 0, 1, \dots, N-1 \quad (4)$$

The x, y is the sampling value for space domain; u, v is the samples values for the frequency domain. In the

digital image processing, digital image usually expressed in pixel phalanx, namely, $M = N$.

III. ALGORITHM PROCESS

We choose an image with black box as the original image, and black border is to ensure that conservation of energy when the original image geometry transform. The original image to remember: $F = \{f(i, j) | f(i, j) \in \mathbb{R}; 1 \leq i \leq N_1, 1 \leq j \leq N_2\}$. $F(i, j)$ is the pixel gray value of the original image. In order to facilitate the operation, we assume that the $N_1 = N_2 = N$.

A. The characteristics of the original image vector selection method

First of all, we put the original image to global DCT transform, and get the DCT coefficient matrix $F_D(i, j)$. Then take in the DCT coefficient matrix of $F(1, 1) \sim F(1, 10)$ ten low intermediate frequency coefficient. We found that when the image is under the geometric attacks, the size of the low part of the DCT intermediate frequency coefficient changed but the symbols (on behalf of the component or 180° in the direction of the phase) basically didn't change. We use "1" to express the positive DCT coefficients, and use "0" to express the negative coefficient (including the zero), as shown in table 1 coefficient of symbol sequence, we observed the column can be found that regardless of conventional attacks or geometric attacks, the symbol can maintain similar sequences and the original image sequence of symbols, and the original symbol sequence correlation coefficient is larger (see last column here took 10 DCT coefficient symbol)

TABLE 1 FULL ORIGINAL IMAGE DCT TRANSFORM COEFFICIENT OF LOW FREQUENCY PART AND DIFFERENT ATTACKS AFTER THE CHANGE OF THE VALUE

	Image manipulation	Part of the coefficient					Symbol sequence	NC
		F(1,1)	F(1,2)	F(1,3)	F(1,4)	F(1,5)		
	original image	96.36	-1.181	-8.795	1.071	0.654	1001100100	1.0
Conventional attack	Gaussian noise	101.4	-1.101	-7.299	0.984	1.365	1001100100	1.0
	JPEG compression	96.93	-1.173	-8.814	1.068	0.863	1001100100	1.0
	median filtering	96.31	-1.183	-8.767	1.073	0.685	1001100100	1.0
Geometric attacks	level movement	97.46	-0.834	-9.909	1.347	1.542	1001100100	0.94
	vertical movement	96.55	-1.051	-7.995	1.271	0.744	1001100100	0.93
	rotating	96.35	-1.215	-9.526	1.115	0.739	1001100100	1.0
	scaling	48.09	-0.576	-4.557	0.547	0.335	1001100100	0.94

*DCT transform coefficient unit $1.0e+003$

For different original image get coefficients of image sequences through the above mentioned methods, and each image is obtained by symbolic operation sequence of the low intermediate frequency DCT symbols (taking the former 32-bit), and characteristic vector, as shown in table 2 indicate different between the original image, feature vector is large, less relevant, and less than 0.5.

TABLE 2 DIFFERENT ORIGINAL IMAGE FEATURE VECTOR
CORRELATION COEFFICIENT (WITH NO BLACK BOX)

	P1	P2	P3	P4	P5
P1	1.00	-0.06	0	0.12	-0.38
P2	-0.06	1.00	-0.31	0.05	-0.05
P3	0	-0.31	1.00	0	-0.38
P4	0.13	0.05	0	1.00	0.03
P5	-0.38	-0.05	-0.38	0.03	1.00

As a result, the DCT coefficient of image symbol sequence can be used as its visual feature vector.

B. Establishing database of image features

Step1: An image feature vector obtained by the DCT transforms.

We let each of the original image in the global DCT transform, and get the DCT coefficient matrix $FD(i, j)$, then to Zig - Zag scanning of DCT coefficient matrix, from low to high frequency DCT coefficients obtained sequence $Y(j)$, and then choose the former L value, and the visual characteristics of the original image is obtained by symbolic operation vector $V(n)$, L is the number of transform coefficient, this paper is 32.

$$FD(i, j) = DCT2(F(i, j)) \quad (5)$$

$$Y(j) = Zig - Zag(FD(i, j)) \quad (6)$$

$$V(n) = Sign(Y(j)) \quad (7)$$

Step2: Putting the original image feature vector in the original image feature database

C. To implement image retrieval

Step3: To obtain the retrieve image feature vector by DCT transform.

Set to retrieve images to $F'(i, j)$, through global DCT transform, get the DCT coefficient matrix $FD'(i, j)$, according to the above Step1, seeks to retrieve image visual feature vector V' .

$$FD'(i, j) = DCT2(F'(i, j)) \quad (8)$$

$$Y'(j) = Zig - Zag(FD'(i, j)) \quad (9)$$

$$V'(n) = Sign(Y'(j)) \quad (10)$$

Step4: Use the peak signal-to-noise ratio (PSNR) evaluation to retrieve image quality after the attack.

Peak signal to noise ratio PSNR of the formula is:

$$PSNR = 10 \lg \left[\frac{MN \max_{i, j} (I(i, j))^2}{\sum_i \sum_j (I(i, j) - I'(i, j))^2} \right] \quad (11)$$

$I(i, j)$, $I'(i, j)$ are the original image and retrieve images. In (i, j) this pixel values, M and N represent the image of the row and column, for the convenience of operation, usually digital image in pixels square, namely, $M = N$. Peak signal to noise ratio is a power signal is possible and destructive affected the accuracy of his said noise power ratio engineering terms, usually adopt peak signal-to-noise ratio as an objective evaluation standard of image quality.

Step5: Matching all visual feature vector $V(n)$ in the database with retrieve image visual feature vector V' , and get the normalized correlation coefficient $NC(n)$.

The normalized correlation coefficient formula $NC(n)$ is :

$$NC(n) = \frac{V(n)V'}{V^2(n)} \quad (12)$$

Step6: Put the picture of the $NC(n)$ is greater than 0.5 presses from big to small order, and supply the user to select pictures.

Fig .1 is the flow chart of the algorithm.

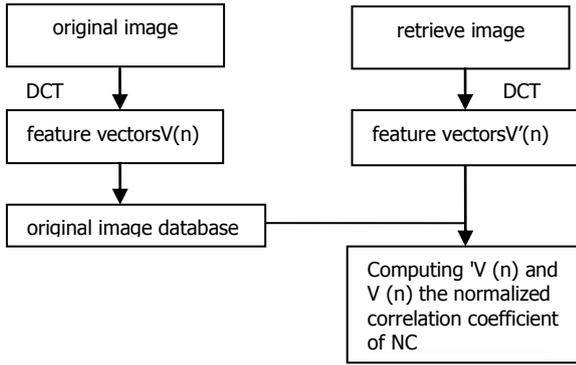


Figure 1 flow chart of the image retrieval algorithm based on DCT transform domain

IV. THE EXPERIMENTAL RESULTS

We are using Matlab2010a simulation platform. The original image is shown in Fig .2 (a), the original image is expressed as $F(i, j)$, $1 \leq i \leq 128, 1 \leq j \leq 128$; The DCT transform coefficient matrix is the $FD(i, j)$, $1 \leq i \leq 128, 1 \leq j \leq 128$. Through the coefficient of low frequency part of its symbolic operation to get the image visual feature vector. Considering the robustness and the time complexity of the algorithm, we take the former 32 coefficient in the low intermediate frequency (a plural has a real part and imaginary part of two coefficient). We judge the retrieve images which is user wanted by calculating the normalized correlation coefficient (NC)

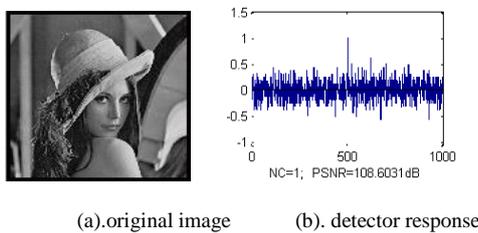


Figure 2 No attack on the similarity of the original image detection

Fig . 2(b) for no attack on the similarity of the original image detection, we can see the $NC = 1.0$, you can retrieve the original image.

A. Conventional attack

1) Gaussian noise

Label to the original texture image with Gaussian noise interference, noise intensity was 20%. At this point, the image PSNR = 9.95 dB, by observing the Fig .3 (a) can be found that images have become blurred; Can see

B. Geometric attacks

3) Rotation transformation

from Fig .3 (b), can still detect the label for the original texture image, $NC = 1.0$.

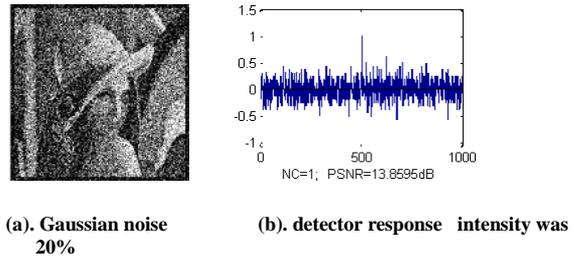


Figure 3 Gaussian noise intensity of 20% retrieval image and similarity detection

TABLE 3 GAUSSIAN NOISE INTERFERENCE EXPERIMENT DATA

noise (%)	1	5	10	20	40	70	90
PSNR(dB)	20.27	13.86	11.48	9.49	7.99	7.13	6.81
NC	1.0	1.0	1.0	1.0	1.0	1.0	1.0

To observe the experimental data can be seen in table 3, when Gaussian noise intensity is as high as 90%, $NC = 1.0$, still can pass the test judging is the original image. It indicates that the algorithm has a good ability to resist Gaussian noise interference.

2) JPEG compression

Original image to JPEG compression, when the compression quality is 4%, the image appears square effect, as shown in Fig .4 (a). See from Fig .4 (b) still can retrieve the original image, $NC = 1.0$.

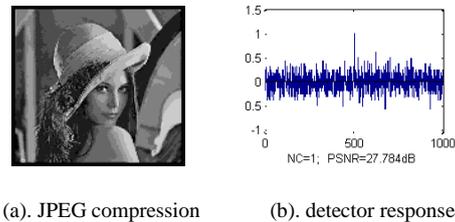


Figure 4 JPEG compression quality of 4% for detecting and similarity retrieval images

TABLE 4 JPEG COMPRESSION EXPERIMENT DATA

quality(%)	2	4	8	10	20	40
PSNR(dB)	25.87	27.78	31.49	32.62	36.05	39.25
NC	1.0	1.0	1.0	1.0	1.0	1.0

Observe the experimental data in table 4, when the compression quality is 4%, with $NC = 1.0$, still can accurately retrieve the original image. As a result, the algorithm has a very ideal JPEG attack resistance.

The original image is 3 degrees clockwise, as shown in Fig .5 (a), Fig .5 (b) as the similarity of detector response, this time the sex ratio of the image is very low,

PSNR = 15.61dB, but NC = 0.76, still can accurately detect for the original image.

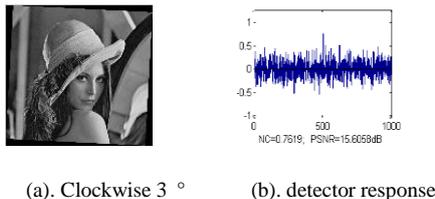


Figure 5 5 °clockwise to retrieve images and similarity detection

TABLE V ROTATION TRANSFORM TO THE EXPERIMENTAL DATA

degree(clockwise)	1°	3°	5°	7°	9°
PSNR(dB)	20.23	15.61	13.86	12.79	12.13
NC	0.94	0.76	0.69	0.63	0.58

By the experimental data in table 5, we observed rotation degrees up to 9 degrees, NC = 0.58, still can be more accurate to detect for the original image. Therefore, the presented algorithm has stronger ability to resist rotation attack.

4) The scaling attack

To reduce 0.5 times as the original image, clear image decreases a lot, as shown in Fig .6 (a). From Fig .6 (b) you can get, NC = 1.00, can be accurately detected for the original image.

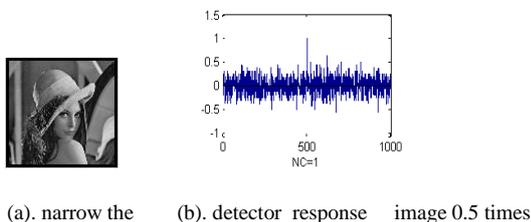


Figure 6 Images of the scaling factors of 0.5 and similarity detection

TABLE VI SCALING ATTACK EXPERIMENTAL DATA

factor	0.4	0.5	0.8	1.2	2.0	4.0
NC	1.00	1.00	1.00	1.00	1.00	1.00

According to table 6 experimental data shows that when the scaling factor is 4.0 can be detected as the original image, NC = 1.00. Therefore, this algorithm has strong ability of anti scaling attack.

V. CONCLUSION

This paper puts forward a kind of image retrieval algorithm based on transform domain. It combines visual feature vector, symbolic operation and database technology. The experiment results show that the algorithm can resist some regular attack and geometric attack ability, and has strong robustness. In addition, the algorithm is realized in the database storage features vector instead of the original image, greatly reduces the storage space. Therefore, this method has good practicality in the process of practical application.

ACKNOWLEDGEMENT

This work is supported by the National Natural Science Foundation of China (No:61263033) and the Institutions of Higher Learning Scientific Research Special project of Hainan Province (Hnkyzx2014-2) and the Key Science and Technology Project of Hainan (No: ZDXM20130078).

REFERENCE

- [1] eklund Yi. The combination of SIFT features and PGH image retrieval method research [D]. Chongqing university, 2013.
- [2] Sun Jun. Content-based image retrieval technology research [D]. Xi'an university of electronic science and technology, 2005.
- [3] Xiang-lin Huang, Zhao zhongxu. Content-based image retrieval technology research [J]. Journal of electronics, cut flower production potentials of 2002:1065-1071.
- [4] Flickner M., Sawhney H., Niblaek W., et al., Query by image and video content: The QBIC System. IEEE Computer, 1995, 28(9): 23 — 32.
- [5] Smith J.R., Chnag5.F., VisualSEEK: a fully automated content-based image query system. ACM Multimedia Nov. 1996, PP. 87 — 98.
- [6] Cllad C., Segre B., Hayit., et al. Bloekworid: image segmentation using Expectation-maximization and its application to image querying. IEEE Trans on PAMI, 2002, 24(8): 1026 — 1038.
- [7] Xu Qing, Yang Weiwei, Chen Shengtang. Content-based image retrieval technology [J]. Computer technology and development, 2008, 01: + 131 126-128.
- [8] Cox I, Kilian J, Leighton T, Shamoon T, "Secure spread spectrum watermarking for multimedia," IEEE Transactions on Image Processing, 6(12), pp. 1673-1687, 1997.
- [9] Ester Yen and Li-Hsien Lin, "Rubik's cube watermark technology for grayscale images", Vol 37(6), pp 4033-4039, Jun. 2010.
- [10] Alar Kuusik, Enar Reilent, Ivor Loobas, Marko Parve, "Software Architecture for Modern Telehealth Care Systems", AISS, Vol. 3, No. 2, pp. 141 ~ 151, 2011.