An Aquatic Organism Image Retrieval Hash Algorithm Based on DCT

Yuanhui Yu[1], Ying Deng[2, a], Yanfen Lin[2]

[1]Computer Engineering College of JiMei University, 361021 Xiamen, China
[2]Xiamen Institute of Technology, 361021 Xiamen, China

Abstract. In order to overcome the disadvantages of traditional image retrieval algorithms such as inefficiency and time-consuming, an image retrieval algorithm based on DCT Hash is proposed. In this paper, the discrete cosine transform technique is applied to image retrieval. Firstly, the feature hash sequence is generated by computing the DCT transform component of the image based on annular segmentation. Then an image Hash algorithm based on DCT feature Hash sequence is proposed. Experimental results show that the algorithm is robust to image scaling, rotation, noise processing, and has good retrieval performance.

1 Introduction

In 2010, Ministry of Science and Technology of China has established "the national natural science and technology resources platform", the platform integrates a variety of aquatic organism resources, and most of these resources are in the form of images. But there are many different kinds of aquatic organism, and their structure is complex, so the retrieval efficiency is inefficient. Therefore, study of advanced aquatic organism image retrieval method, to improve the accuracy and efficiency of image retrieval and share the aquatic organism resources better, has become an urgent problem to be solved, this is also the important component of the optimization of aquatic organism resources platform of science and technology.

At present, the key of traditional content-based image retrieval technology is to extract the texture, colour, shape and other visual feature vectors to compare the similarity. However, because the images are usually massive, and the feature vectors of the images are also high-dimensional, the efficiency of image retrieval is seriously affected [1]. Image Hash is a new research topic in the field of multimedia information security. This technique converts image feature data into binary coding sequences of dozens or hundreds of bits by Hash function, which is called Hash coding. The storage space needed to represent image data by hash coding will be greatly reduced and the retrieval speed will be improved. Due to the excellent performance of Hash algorithm in high dimensional data search, the Hash based data search method has become a hot topic in the research of image retrieval, and many classical Hash algorithms have been proposed, such as hash algorithms based on discrete Fourier transform (DFT) domain [2, 3] discrete wavelet transform (DWT) domain [4]], discrete cosine transform

(DCT) domain [5], integral transform Radon domain [6], etc. However, there are still many problems to be solved in practical applications. For example, the performance of uniqueness and rotation robustness is not ideal [7]. To solve this problem, this paper proposes an image Hash algorithm based on annular segmentation of discrete cosine transform (DCT) feature points. The algorithm first pre-processes the input image to generate a normalized image. On this basis, the normalized image is segmented in a circle and a two-dimensional DCT transform is performed. The DC component of the DCT is used to construct the feature vector to generate the image Hash sequence. Experimental results show that the proposed algorithm is robust to normal digital processing and has good uniqueness.

2 DCT Hash algorithm based on annular segmentation

Texture feature is the common inherent property of all objects' surface, containing the important information of an object's surface structure arrangement and their relations with the surrounding environment. Structure analysis and statistical analysis method are two main methods of texture analysis. Structural analysis methods include morphological operator, boundary figure and so on. Structure analysis method is very effective for regular texture image. Statistical analysis methods include: co-occurrence matrix, Tamura texture feature, Markov random field, fractal model, Gabor and other multi-resolution analysis methods, and wavelet transform etc. [1] Discrete cosine transform is an orthogonal transform based on real number, and it is widely used in various kinds of video and image coding standard due to its higher compression performance and rapid and simple realizing algorithm. The author mainly

---

[a] Corresponding author: yingd_529@163.com

studies the technique of discrete cosine transform, and applied it to the aquatic organism's image retrieval.

2.1 The basic principle of discrete cosine transform

Discrete cosine transform is proposed by N.Ahmed, T.Natarajan and K.R.Rao in 1974 [8, 9].Two-dimensional DCT transform is often used in digital image processing. For a block of M × N pixels, the two-dimensional discrete cosine transform (DCT) is defined as in Eq. (1):

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

Where u=0, 1… M-1, v=0, 1… N-1.

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

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

x, y are spatial sampling values, u, v are frequency domain sampling values.

In Eq.(1), $F(u,v)$ represents the high-frequency component of the transform domain, also known as the AC coefficient, and F(0,0) represents the low-frequency component in the transform domain, also known as the DC coefficient. F(0,0) is the concentration of all sample values in a block, which represents the mean of all amplitude of the input matrix. With the increase of u and v, the corresponding coefficient respectively representing gradually increasing size of the horizontal spatial frequency component and the vertical spatial frequency component.

2.2 DCT Hash algorithm based on annular segmentation

For a colour RGB aquatic organism image, DCT Hash algorithm based on annular segmentation is described as follows:

(1) Firstly, a colour RGB aquatic organism's image is converted to grayscale image, That is, the luminance component is extracted, and the formula is as follows:

I=0.114×B+0.587×G+0.298 9×R   (4)

(2) The bilinear difference method is used to adjust the size of the processed grayscale image to 256×256.The purpose of this step is to make the generated Hash sequence length consistent.

(3) First, an image is divided into five parts, a small circle and four concentric annuluses. Then the five parts are filled with black to form a square region. As shown in Fig. 1, Fig. 2 and Fig. 3.

(4) Each square is divided into 8 × 8 sub-blocks, and then perform discrete cosine transform on each image block.

(5) For each block after discrete cosine transform, the DC component of its transform coefficient is extracted, and the mean μ of these 64 DC components is calculated. The mean μ as follows:

$$\mu = \frac{1}{64} \sum_{k=1}^{64} I_k \quad (5)$$

Where I_k (k = 1, 2, 3…64) represents the kth sub-block DC component [10].

(6) For each square, 64 sub-blocks are traversed respectively, and the DC component I_k of each sub-block is compared with the mean value. If $I_k \geq \mu$, it is recorded as 1, otherwise, it is recorded as 0, and a binary string $\mathbf{h}_i$ with a length of 64 bits is obtained. As shown in Fig. 4.

(7) Concatenate the hash sequences of each sub-block to form the hash sequence of the entire image, that is

$$\mathbf{h}= [\mathbf{h}(1), \mathbf{h}(2)…, \mathbf{h}(n)] \quad (6)$$

(8) Hash similarity calculation. In this paper, we use normalized Hamming distance to measure the similarity between two Hash sequences h1 and h2:

$$d(\mathbf{h_1}, \mathbf{h_2}) = \frac{1}{L_h} \sum_{i=1}^{n} |\mathbf{h_1}(i) - \mathbf{h_2}(i)| \quad (7)$$

Where L_h is the length of Hash sequence, and $\mathbf{h_1}(i)$ and $\mathbf{h_2}(i)$ are the binary bits in the corresponding Hash sequence respectively. The smaller the distance, the more similar the image is, and the larger the distance, the greater the difference [11].
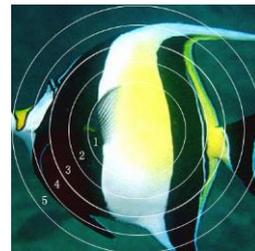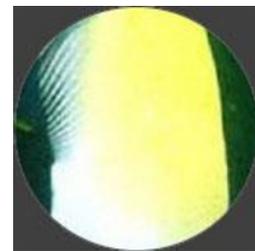

Figure 1. Schematic of annular geometry segmentation

Figure 2. Filling effect of center circle 1.

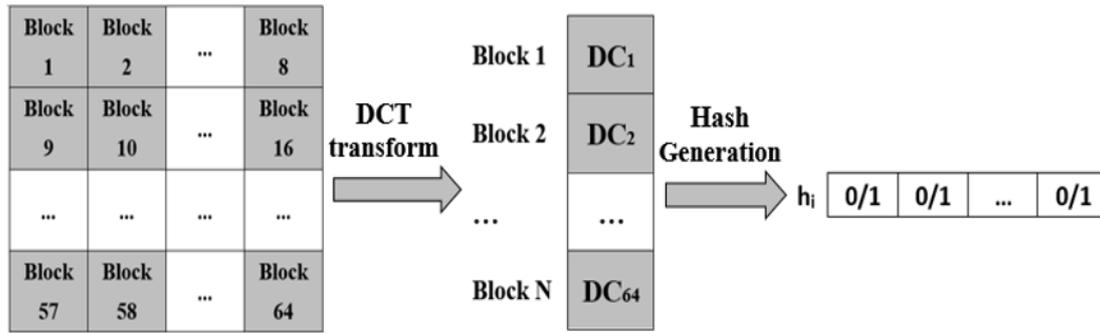**Figure 3.** Filling effect of annulus 2.

Figure 4. Extraction process of image Hash sequence.

3 Experimental results and analysis

3.1 Perceptual robustness

The experiment use a test database consisting of 200 colour images [12], in order to further validate the retrieval method in this paper is robust for the image translation, rotation and scaling, three images were selected at random from the test database, rotate 5, 10,15,45,90 and 180 degrees, add Gaussian noise , zoomed out 2 times, zoomed in 2 times respectively. In the experiment, the size of the image block is 8×8, so the length of the Hash generated by the method in this paper is $8 \times 8 \times 5 = 320b$. Table 1 shows the robust performance of different digital operations. The experimental results show that most of the normalized Hamming distances of all operations are less than 0.1. This shows that the algorithm is robust to image scaling, rotation and noise processing. Moreover, the hamming distance with rotation angle less than 180 degrees is less than 0.15, which shows that the algorithm is robust to rotation.

Table 1. Robust performance of different digital operations.

Operations	Image 1	Image 2	Image 3
Zoomed out 2 times	0	0	0
Zoomed in 2 times	0	0	0
Rotate 5 degrees	0	0	0
Rotate 10 degrees	0.0156	0	0.0156
Rotate 15 degrees	0.0313	0.0156	0.0313
Rotate 45 degrees	0.0625	0.0469	0.0313
Rotate 90 degrees	0.1094	0.0781	0.0781
Rotate 180 degrees	0.1250	0.0469	0.0938
add Gaussian noise	0	0	0

3.2 Uniqueness

Uniqueness is that two images with distinct visual differences should produce different perceptual Hash sequences, that is, the normalization of the Hamming distance between the two visually different images of Hash sequences is greater than a certain threshold T, and otherwise there will be a conflict. Table 2 shows true positive rate and false positive rate under different thresholds. In Table 2, 13.13% of different images are misjudged as the same image when T = 0.4, and 7.82% of different images are misjudged as the same image when T = 0.3. In fact, a small threshold can reduce the misjudgement rate of different images, but increase the misjudgement rate of similar images. The experimental results show that the total false positive rate corresponding to T = 0.3 is 13.578, which is smaller than the total false positive rate of other thresholds, so 0.3 is considered as the best threshold.

Table 2. True positive rate and false positive rate under different thresholds.

Threshold T	True positive rate (%)	False positive rate (%)	Total false positive rate (%)
0.6	0	28.6	28.6
0.5	0.27	15.64	16.18
0.4	1.546	13.13	14.676
0.3	5.938	7.82	13.578
0.2	10.94	4.688	15.628
0.1	19.8	0.5	20.3

3.3 Algorithm performance comparison

The retrieval performance of this algorithm is compared with that of reference [13], in which we use the Precision vs. Recall curves as an evaluation standard of retrieve effect of algorithms. Precision refers to the ratio of the number of relevant images r returned by the system to the number of all returned images N of the system in the process of a query. Recall refers to the ratio of the number of relevant images r returned by the system to the number of all relevant images R (includes the number of returned and unreturned images) in the process of a query. By performing all queries of retrieve

collection, we can calculate the average recall and precision of the query. In general, the curve closest to the top of the chart indicates the best performance [14]. From Fig. 5, we can see that the precision / recall curve of the proposed method is located at the top of the graph, which shows that the method has better retrieval results. Moreover, the algorithm runs faster, the average search time is 2.458s, and the search time of the latter is 7.46S.
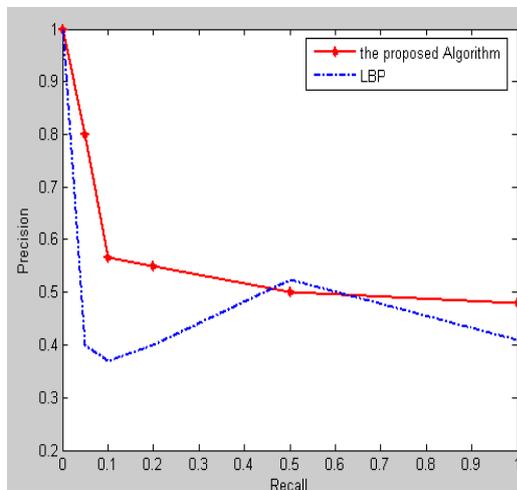

Figure 5. Interpolated precision-recall averages.

4 Conclusion

This paper presents a DCT-based image Hash feature sequence extraction method. The Hash algorithm first transforms the color image into a gray image, and adjusts the image into a unified size by the bilinear interpolation method. Then the ring geometric segmentation is used to extract the DC component of the sub block by DCT transform in different segmentation regions, and the shorter Hash sequence is generated by the feature distance. The algorithm presented in this paper has better robustness and uniqueness for the common operations that keep the image content constant, such as image scaling, rotation and noise processing. The current Hash algorithm only considers the gray image, and the future research work will include the Hash algorithm design, which can extract the image color, shape and other related features to further optimize the algorithm and improve the algorithm's retrieval performance.

Acknowledgments

References

1. YueTing Zhuang, Quhe Pan, Fei Wu, *Online Multimedia Information Analysis and Retrieval, Tsinghua University Press*, 21(2002).

2. A. Swaminathan, Y Mao and M. Wu, Robust and secure image hashing. *IEEE Transactions on Information Forensics and Security*.**1**, 215.

3. M. Wu, A. Swaminathan and Y Mao, A signal processing and randomization perspective of robust and secure image hashing, SSP'07, 166(2007).

4. R.Venkatesan, S. McKeon, M.H.Jakubowski and P. Moulin. Robust image hashing. ICIP 2000. 664(2000).

5. F. Lefebvre, B. Macq, J.D. Leuat. RASH: Radon soft hash algorithm. *Proceedings of llth European Signal Processing Conference*, 299(2002).

6. S.S. Kozat, R. Venkatesan, M.K. Mihcak. Robust perceptual image hashing via matrix invariants. *Proceedings of 2004 International Conference on Image Processin*, 3443(20040).

7. Zhenjun Tang, Fan Yang, Ziqing Huang, Hua Lao. Image Hashing Algorithm Based on PCA Feature Distance. Journal of Guangxi Normal University (Natural Science Edition), **4**, 9(2016).

8. N. Ahmed, T. Natarajan, K.R. Rao., Discrete Cosine Transform. IEEE Transactions on Computers (1974).

9. XiXun Zhang, The design and implementation of the discrete cosine transform, *Ph.D. Thesis, Zhejiang University* (2006).

10. Yuanhui Yu and Ying Deng, The Approach of Image Retrieval Combining Energy Moment and Color Clustering, Journal of Henan University: Natural Science, **1**, 96(2012).

11. OU Yang, RHEE K H. A key-dependent secure image hashing scheme by using Radon transform. *Proceedings of the 2009 international Symposium on intelligent Signal Processing and Communication Systems*, 595(2009).

12. Information on http://image.baidu.com.

13. T. Ojala, M. Pietikäinen, T. Mäenpää, Multiresolution Gray-Scale and Rotation Invariant Texture Classification with Local Binary Pattern, *IEEE Trans. on PAMI*, **24**,971(2002).

14. R.O. Stehling, M.A. Nascimento, A.X. Falcao, A Compact and Efficient Image Retrieval Approach Based on Border/Interior Pixel Classification, *Proceedings of the eleventh international conference on Information and knowledge management*, 102(2002).