Block Mean Value Based Image Perceptual Hashing

Bian Yang[1], Fan Gu[2] and Xiamu Niu[1,3]
[1]*School of Computer Science and Technology, Harbin Institute of Technology, China;*
[2]*Center for Coastal and Ocean Mapping, University of New Hampsire, U.S.A.*
[3]*Shenzhen Graduate School, Harbin Institute of Technology, Shenzhen, China*
xiamu.niu@ict.hit.edu.cn

Abstract

Image perceptual hashing has been proposed to identify or authenticate image contents in a robust way against distortions caused by compression, noise, common signal processing and geometrical modifications, while still holding a good discriminability for different ones in sense of human perception. We propose and compare four normalized block mean value based image perceptual hashing algorithms which demonstrate higher performances than other existing algorithms in robustness-and-discriminalibility and simplicity for implementation. Overlapped blocking and rotation operations are employed to enhance the robustness to geometrical distortions. To evaluate the proposed algorithms' robustness and discriminability, given fixed modifications, identification ratio is used; and given fixed content classification, receiver operating curves is obtained.

1. Introduction

The reliable identification or authentication of content is popular in the process of storing and distribution of digital information. Cryptographic hashing techniques digest input data into short binary strings and have been used for identification of binary data like executables. However, most cryptographic hashes are sensitive and meet difficulties when coming to multimedia content which has different representation versions, such as format conversion and compression. Thus, the technology called perceptual hash is proposed to identify or authenticate the content of digital media with various representation versions, and only those manipulations that change the content apparently could affect the calculated perceptual hash function. Technologies for the identification of multimedia content are also included in the ongoing standardization activities such as MPEG-7 or MPEG-21, or as the so called 'Persistent Association Technologies' (PAT) [1].

Considering its application scenarios, as stated in [2], perceptual hashing should satisfy several requirements including ability of discrimination, robustness against attacks, conciseness, computational complexity of hash calculation and retrieval, algorithm security, etc.

A general scheme of the perceptual hashing technology is described in [3]. Normally, image features invariant to allowed content-preserving image processing operations are identified and used to generate the hash functions. Image edge information [4] was initially proposed to generate the hash function, but the resulting hash could not be tolerant to rotation distortion. Seo et al. [5] proposed an affine transform resilient image fingerprinting algorithm based on the auto-correlation of the Radon transform, the log mapping and the Fourier transform was used to deal with affine transform. Venkatesan et al. [6] presented a technique to divide a wavelet transformed image into tiles and extract the statistics of tiles as hashes. Fridrich [7] proposed a robust hashing method whose hashing vector were created by projections of DCT coefficients to key-dependent random patterns. Lu et al. [11] proposed a robust mesh-based hashing to resist more geometrical distortions. Yu et al. [8] stated a cumulant-based image fingerprint. In general, few of these algorithms have good performance in both robustness-discrimination and computational complexity.

In this paper, we introduce and compare four normalized block mean based image perceptual hashing functions. The first two methods are based on normalized block mean values while the third and forth methods incorporate a rotation operation, which enhances robustness against the rotation attacks. At the same time, the second and forth methods are overlapped version of the first and third methods, respectively. We discuss the proposed algorithms' advantages compared with other algorithms in section 4. Security is achieved by encrypting the hashing values by secret key in all proposed algorithms.

2. Block mean value based algorithms

2.1. Method 1

The first perceptual hashing function is based on the mean values of the blocks, described as follows:

a) Normalize the original image into a preset sizes;
b) Divide the size-normalized image I into non-overlapped blocks $I_1, I_2, \ldots, I_N$, in which N is the

Proceedings of the 2006 International Conference on Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP'06)
0-7695-2745-0/06 $20.00 © 2006 IEEE

block number equal to length of the final hash bit string;

c) Encrypt the indices of the block sequence $\{I_1, I_2, \ldots, I_N\}$ using a secret key K to obtain a block sequence with a new scanning order $\{I'_1, I'_2, \ldots, I'_N\}$;

d) Calculate the mean value sequence $\{M_1, M_2, \ldots, M_N\}$ from corresponding block sequence $\{I'_1, I'_2, \ldots, I'_N\}$ and obtain the median value M_d of this sequence as:

$$M_d = \text{median}(Mi) \qquad (i=1,2,\ldots,N) \tag{1}$$

e) Normalize the mean value sequence into a binary form and obtain the hash values h as:

$$h(i) = \begin{cases} 0, M_i < M_d \\ 1, M_i \geq M_d \end{cases} \tag{2}$$

2.2. Method 2

The second perceptual hashing function takes advantage of overlapping. The degree of overlapping is set to be half-size of a block and Method 1 is taken to calculate hash values.

2.3. Method 3

To have higher robustness against the rotation attack, the third perceptual hashing function incorporates rotation based on theception hashing function. Steps are as follows:

a) Normalize the original image into a preset sizes;

b) Divide the size-normalized image I into unoverlapped blocks $I_1, I_2, \ldots, I_N$, in which N is the block number equal to length of the final hash bit string;

c) Encrypt the indices of the block sequence $\{I_1, I_2, \ldots, I_N\}$ using a secret key K to obtain a block sequence with a new scanning order $\{I'1, I'2, \ldots, I'_N\}$;

d) Calculate the mean value sequence $\{M_1, M_2, \ldots, M_N\}$ from corresponding block sequence $\{I'_1, I'_2, \ldots, I'_N\}$ and obtain the median value M_d of this sequence as in Method 1;

e) Rotate by D degrees the matrix M formed by $\{M_1, M_2, \ldots, M_N\}$, $D = \{0, 15, 30, \ldots, 345\}$. Divide rotated matrix M_i (i = 1, 2, …, 24) into N blocks. Obtain the mean value sequence $\{M_{i1}, M_{i2}, \ldots, M_{iN}\}$ of each block and median value M_{di} of this sequence, which forms 24 groups of sequences;

f) Perform the same equation (2) for the 24 groups respectively and obtain the final hash value matrix.

2.4. Method 4

Method 4 is also based on the rotated mean values of the blocks (Method 3) but incorporates an overlapping operation same as in Method 2.

3. Test Metrics

We chose the following three metrics to characterize the robustness, discriminability of the perceptual hashing function: bit error rate (BER), hit rate, and receiver operation characteristics (ROC).

3.1. Bit error rate (BER)

Considering the binary form of perceptual hash, the number of mismatched bits i normalized by the number of bits of the perceptual hash n describes the distance between two perceptual hashes. It is called bit error rate (BER) as

$$\rho = \frac{i}{n} \tag{3}$$

where $i \in \{0, 1, 2, \ldots, n\}$ and $0 \leq \rho \leq 1$. The smaller is BER, the higher is the possibility that the corresponding multimedia data contains perceptually identical content.

3.2. Hit rate

In our case, it describes the accuracy of detection, which is defined by the rate of right detections over all original input contents for a certain attack or modification.

3.3. Receiver operating characteristics (ROC)

In most applications, the process of image authentication is similar to a hypothesis testing process with the following two hypothesizes [9]:

H1: Image is not authentic;

H2: Image is authentic;

The ROC curve demonstrates the receiver's performance by classifying the received signal into one of the hypothesis states. It is built upon false acceptance rate (FAR), which is the probability that hypothesis H2 is accepted when the hypothesis H1 is true; and false rejection rate (FRR), which is the probability that the hypothesis H1 is accepted when the hypothesis H2 is true. ROC curve is a plot of the probability of detection 1-FAR for the false alarm probability FRR. The closer the curve is to the corner with small FRR value and large 1-FAR value, the higher detection ability the algorithm has.

4. Experimental results

Our experiment is performed on a database of 2600 images with 100 original gray scale images consisting of classic benchmark images, such as Lena, Baboon, Pepper, etc., and a variety of photos from [12]. These photos were converted to gray scales and normalized to 256×256. For each original image, 25 manipulation versions are generated with the distortions as follows:

Compression: 1.JPG Q=10; 2.JPG Q=40; 3.JPG Q=90;
Filtering Operation: 4.Gaussian 3×3 filtering; 5.Sharpening 3×3 enhancement; 6.Median filtering 3×3;
Geometric Distortions: 7.Scaling 0.5; 8.Scaling 2.0; 9.Cropping 2%; 10.Cropping 10%; 11.Cropping 20%; 12.Random removal of 5%; 13.Shearing 2%; 14.Shearing 10%; 15.Rotation 2 degree + scaling + cropping; 16.Rotation 5 degree + scaling + cropping; 17.Rotation 10

degree + scaling + cropping; 18.Rotation 15 degree + scaling + cropping; 19.Rotation 30 degree + scaling + cropping; 20.Rotation 90 degree + scaling + cropping
Additive Noise: 21.Gaussian noise 0.01; 22.Salt&Pepper noise 0.02; 23.Speckle noise 0.01
Luminance Distortion: 24.Histogram Equalization; 25.Brightness Enhancement

4.1 Bit error rate results

Bit error rate for the first two methods are presented in Table 1, consisting of two parts: different contents and identical contents. Distriminability is demonstrated by the BER of different content, while robustness shown by that of the same content with different attacks. After overlapping, robustness is enhanced except for shearing, cropping and rotation. Compared with results from [5] and [8], the proposed two methods have better performance except for larger degree rotations due to the blocking process.

4.2 Hit rate results

Hit Rate is calculated and illustrated in Fig.1. It can be seen that most of the results in Method 2 are higher than Method 1, while most of the values of Method 4 are higher than Method 3. In the rotation modifications, Method 3 and 4 have higher robustness.

4.3 ROC results

ROC curves were plotted in Fig.2-Fig.4. Each figure demonstrates the discriminability against a specified modification. Taking the remaining 25 versions relevant to the input image as the same content, and other original images as different contents, we obtain two Gaussian-shape distributions. Through threshold adjusting, we can obtain a ROC curve for a specified modification. In general, the proposed four methods show desirable discrimination between different contents and the ROC curves are closer to the left upper corner, which bears higher 1-FAR values and lower FRR values. Here we only present several cases whose performances are not so satisfactory. It is obvious the third and fourth methods with rotation operations are better than the former two methods especially for severe modifications such as sharpening, large degree rotation+scaling+cropping, and histogram equalization.

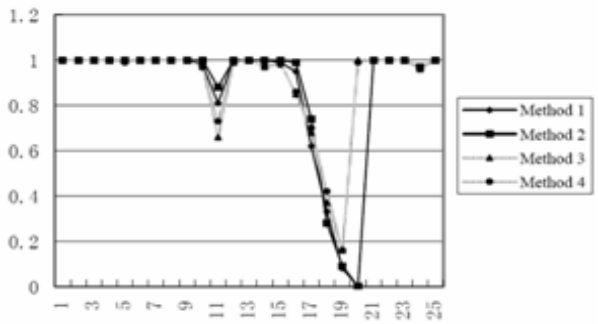

Fig.1. Hit Rate illustration. X-axis: the corresponding modifications and Y-axis: the Hit Rate values

Table 1 Bit Error Rate of the first two method for four example images. The first four rows in the table are the original images and the next 25 rows are 25 modifications of the original images

	Method1				Method2			
Images and Modifications	Lena	Baboon	Barbara	Airplane	Lena	Baboon	Barbara	Airplane
Lena	0.000	0.500	0.550	0.461	0.000	0.476	0.500	0.386
Baboon	0.496	0.000	0.531	0.457	0.477	0.000	0.476	0.410
Barbara	0.551	0.531	0.000	0.480	0.500	0.476	0.000	0.433
Airplane	0.461	0.457	0.480	0.000	0.387	0.410	0.433	0.000
JPG Q=10	0.004	0.008	0.008	0.004	0.004	0.012	0.008	0.004
JPG Q=40	0.004	0.000	0.004	0.000	0.004	0.000	0.000	0.004
JPG Q=90	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Gaussian 3-by-3 filtering	0.000	0.000	0.004	0.000	0.004	0.000	0.000	0.000
Sharpening 3-by-3	0.012	0.027	0.023	0.004	0.008	0.015	0.031	0.004
Median filtering 3-by-3	0.000	0.004	0.000	0.000	0.000	0.008	0.004	0.008
Scaling 0.5	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Scaling 2.0	0.004	0.000	0.004	0.000	0.012	0.000	0.000	0.004
Cropping 2%	0.012	0.043	0.074	0.027	0.035	0.027	0.035	0.012
Cropping 10%	0.246	0.219	0.234	0.121	0.145	0.125	0.136	0.066
Cropping 20%	0.328	0.340	0.363	0.219	0.273	0.257	0.257	0.125
Rand. removal 5%	0.012	0.020	0.027	0.016	0.035	0.031	0.027	0.031
Shearing 2%	0.016	0.059	0.070	0.020	0.039	0.015	0.039	0.012
Shearing 10%	0.211	0.141	0.218	0.070	0.156	0.113	0.105	0.039
Rotation 2 degree + scaling + cropping	0.070	0.098	0.066	0.051	0.102	0.035	0.058	0.054
Rotation 5 degree + scaling + cropping	0.219	0.168	0.089	0.125	0.160	0.136	0.136	0.093
Rotation 10 degree + scaling + cropping	0.313	0.273	0.226	0.215	0.227	0.230	0.222	0.160
Rotation 15 degree + scaling + cropping	0.375	0.352	0.328	0.246	0.301	0.305	0.304	0.195
Rotation 30 degree + scaling + cropping	0.426	0.484	0.468	0.395	0.383	0.387	0.382	0.308
Rotation 90 degree + scaling + cropping	0.480	0.500	0.585	0.492	0.426	0.430	0.468	0.414
Gaussian noise 0.01	0.000	0.004	0.004	0.004	0.012	0.000	0.008	0.008
S&P noise 0.02	0.004	0.000	0.004	0.000	0.004	0.000	0.008	0.004
Speckle noise 0.01	0.000	0.004	0.004	0.004	0.000	0.012	0.004	0.000
Histogram Equalization	0.008	0.008	0.012	0.188	0.070	0.008	0.016	0.137
Brightness Enhancement	0.012	0.012	0.023	0.008	0.035	0.016	0.031	0.004

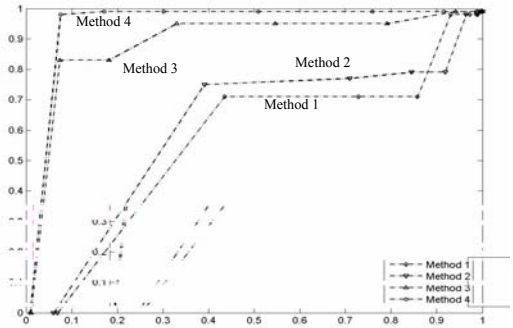

Fig.2. ROC curve of four methods for the 5th modification: Sharpening 3-by-3 enhancement. X-axis is the FRR values, and the Y-axis is 1-FAR values

Proceedings of the 2006 International Conference on Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP'06)
0-7695-2745-0/06 $20.00 © 2006 IEEE

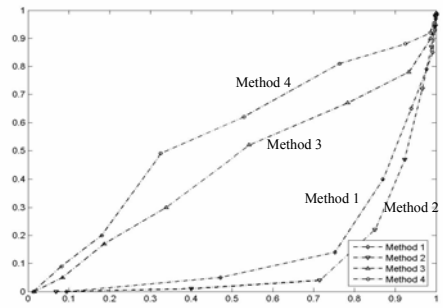

Fig.3. ROC curve of four methods for the 16th modification: Rotation 15 degree + scaling + cropping. X-axis is the FRR values, and the Y-axis is 1-FAR values

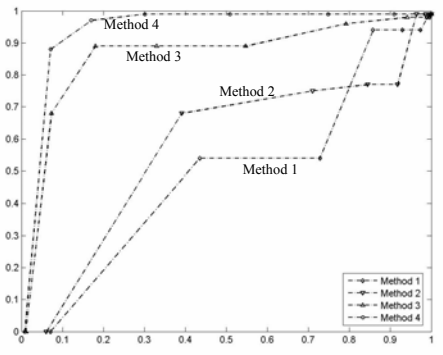

Fig.4. ROC curve of four methods for the 24th modification: Histogram Equalization. X-axis is the FRR values, and the Y-axis is 1-FAR values.

5. Conclusions

Four normalized block mean values based hashing functions were proposed and experimental results were compared. The overall performance of the four methods is better than the previous methods in literatures in terms of robustness, discriminability and computational complexity. From the results it can be concluded that there is a desirable trade-off between the robustness and discriminability from our proposed methods for lossy compression, most common signal processing modifications and rotations with small and middle degrees (<10 degree). Another advantage of the proposed methods is the low computational complexity except the latter two methods employing rotation operations. Our future work focuses on improvements of robustness to large degree geometrical distortions while keeping block-wise methods' efficiency.

6. Acknowledgement

This work was supported by the National Natural Science Foundation of China (Project Number: 60372052, 60671064), the Science Foundation of Guangdong Province (Project Number: 05109511), the Foundation for the Author of National Excellent Doctoral Dissertation of China (Project Number: FANEDD-200238), the Foundation for the Excellent Youth of Heilongjiang Province, the Program for New Century Excellent Talents in University (Project Number:NCET-04-0330), and the Chinese national 863-Program (Project Number:2005AA733120).

7. References

[1]E. N. of Excellence, "Applications, Application Requirements and Metrics", *Technical report*, ECRYPT, 2006.

[2]E. N. of Excellence, "First summary report on forensic tracking", *Technical report*, ECRYPT, 2005.

[3]P. Cano, E. Baltle, T. Kalker, and J. Haitsma, "A review of algorithms for audio fingerprinting", *IEEE Workshop on Multimedia Signal Processing*, pp. 169-73, 2002.

[4]M. P. Queluz, and Twards Robust, "Content Based Techniques for Image Authentication", *IEEE Signal Processing Society-Second Workshop on Multimedia Signal Processing*, pp. 297-302, 1998.

[5]S.J.S., J. Haitsma, T. Kalker, and C. Yoo, "Affine transform resilient image fingerprinting", ICASSP'03, *Proc. IEEE 3*, pp. 61-64, 2003.

[6]R. Venkatesan, S. M. Koon, M. H. Jakubowski, and P. Moulin, "Robust Image Hashing", *Proc, IEEE ICIP 2000,* Vancouver, CA , 2000.

[7]J. Fridrich and M. Golian, "Robust Hash function for Digital Watermarking", *IEEE Proc. International Conference on Infromation Technology: Coding and Computing*, 2000.

[8]L. Yu, M. Schmucker, C. Busch, and S. Sun, "Cumulant-based image fingerprints. Security", *Steganography, and Watermarking of Multimedia Contents VII, Proc. Of SPIE-IS&T Electronic Imageing SPIE* Vol. 5681, 68-75.

[9]A.Swaminathan, Y. Mao, and M. Wu, "Robust and Secure Image Hashing", *IEEE Tran. on Information Forensics and Security*, 2005.

[10]M. Mihcak and R. Venkatessan, "A Perceptual Audio Hashing Algorithm: A Tool For Robust Audio Identification and Information Hiding", *Proceedings of 4th International Information Hiding Workshop*, Pittsburgh, PA, 2001.

[11]C. Lu, C. Hsu, S. Sun and P. Chang, "Robust Mesh-based Hashing for Copy Detection and Tracing of Images', *Proc. IEEE Int. Conf. on Multimedia and Expo: special session on Media Identification*, Taipei, 2004

[12]J.Li, "Photography Image Database", http://www.start.psu.edu/jiali/index.download.html