A Novel Block-DCT and PCA Based Image Perceptual Hashing Algorithm

Zeng Jie

College of Information Engineering, Shenzhen University
Shenzhen, Guangdong, P.R.China

Abstract

Image perceptual hashing finds applications in content indexing, large-scale image database management, certification and authentication and digital watermarking. We propose a Block-DCT and PCA based image perceptual hash in this article and explore the algorithm in the application of tamper detection. The main idea of the algorithm is to integrate color histogram and DCT coefficients of image blocks as perceptual feature, then to compress perceptual features as inter-feature with PCA, and to threshold to create a robust hash. The robustness and discrimination properties of the proposed algorithm are evaluated in detail. Experimental results show that the proposed image perceptual hash algorithm can effectively address the tamper detection problem with advantageous robustness and discrimination.

Keywords: *image hash; perceptual hash; tamper detection; PCA*.

1. Introduction

Image perceptual hashing, also known as image robust hashing, is defined as mapping images to a short bit string following the human perception [1]. In contrast to classic hash functions (MD5, SHA-1), which is highly sensitive to every bit of input data, image perception hashing is sensitive to image content rather than the integrity of image date. The two principal properties of image perception hashing are robustness and discrimination. Robustness means that the hash algorithm should result in the same out bit string for images with the same underlying content. For example, the raw image, its added noise version, its compressed version, its changed brightness version and its rotation angle version have the same underlying content and should share the same hash value. Discrimination implies that the hash values for any two distinct images should be different and random. That is to say, image perceptual hash functions are statistically independent to different image content.

Image perceptual hash value can be used for content identification and digital signature. The former is mainly used in content indexing and analysis, large-scale image database management. The latter is mainly used in the image certification and authentication, digital watermarking. According to the needs of applications, image perceptual hashing should also meet other two properties — randomness and scale-independence. Randomness means that the hash function should withstand all kinds of forgery attack since the hash values are impossible to be reconstructed by the attacker. Scale-independence implies that the length of hash values should always be an even number, although the input images are in different resolution.

Many image perceptual hash functions have been proposed in the literature. Bian Yang[2] uses the mean of image blocks to obtain a perceptual hash. J.Fridrich[3] extracts perceptual features by projecting image blocks onto key based random patterns and thresholds to create a robust hash with the median. R.Venkatesan[4] and M.K. Mihcak[5] selects the low-frequency sub-band of wavelet coefficients to generate a perceptual hash. F.Lefbvre[6] and J.S.Seo[7] uses radon transform to produce a perceptual hash. Hui Zhang[8] creatively introduces the human visual system to obtain a image perceptual hash.

This article addresses the problem of the tamper detection problem of images with image perceptual hashing. Although there are so many image perceptual hash methods proposed, the tradeoff between robustness and discrimination is relatively few discussed. In this article, we aim at proposing a robust and discriminative image perceptual hash algorithm, and explore the algorithm in the application of tamper detection.

The rest of this paper is organized as follows. Section Ⅱ describes the proposed robust and discriminative image perceptual hash algorithm. The experimental results are detailed in Section Ⅲ. Section Ⅳ contains the properties of the hash value. Conclusion and future work are introduced in Section Ⅴ.

2. Robust and Discriminative Image Perceptual Hashing

The main idea of the algorithm is to integrate color histogram and low-frequency Discrete Cosine Transform (DCT) coefficients of image blocks as perceptual features, then to compress perceptual features as inter-features with

IJCSI International Journal of Computer Science Issues, Vol. 10, Issue 1, No 3, January 2013
ISSN (Print): 1694-0784 | ISSN (Online): 1694-0814
www.IJCSI.org

Principal Component Analysis (PCA), and to threshold to create a robust hash. The framework of this algorithm is shown in Figure 1.

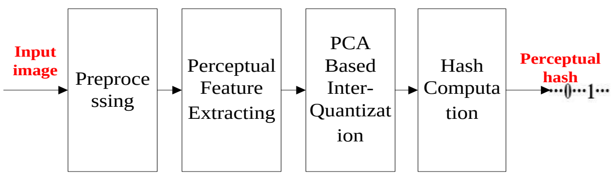


Figure 1. Flow chart of our image perceptual hash method

2.1 Image Preprocessing

The research of cognitive psychology and human visual system show that the sensitivity of eyes to chroma signal is much weaker than to luminance signal, and that brightness is the main features of the image signal[1]. So, only the luminance information is considered in preprocessing. The input image is first converted to a standardized image (64*64) via resampling and interpolation.

Preprocessing not only reduces the computational complexity of follow-up steps (perceptual feature extracting, PCA based inter-quantization, hash computation), but also ensures that the algorithm is independent of scale.

2.2 Perceptual Feature Extracting

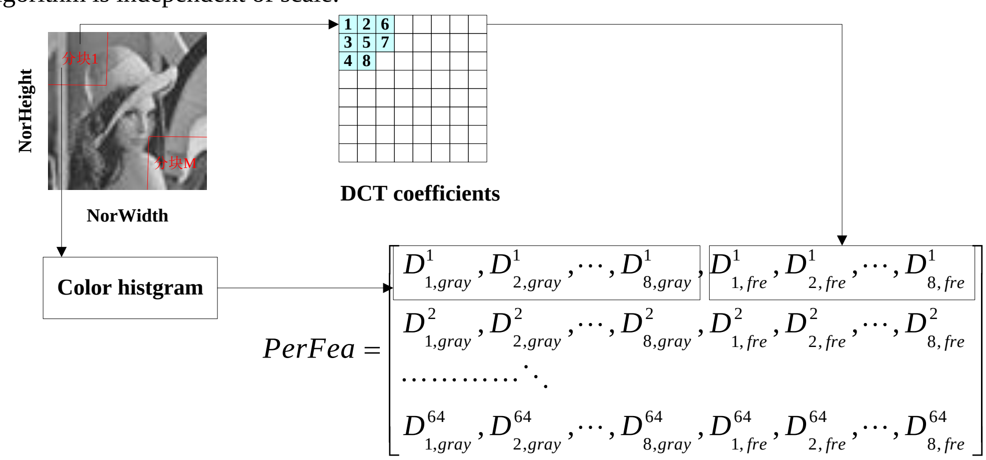
During perceptual feature extracting, we adopt the block images strategy, and integrate color histogram and low-frequency Discrete Cosine Transform coefficients of every image block as perceptual features. The process is detailed as follows:

- Divide the standardized image into 64 blocks (block size: 8*8).
- Calculate the color histogram of blocks successively, and the calculation formula is as follows:

$$hist(i) = count(|ima_{gray} / 32|), \forall i \in 0, 1, \cdots, 7 \qquad (1)$$

- Select the DCT coefficients (DC coefficient and 7 AC coefficients) of blocks successively, and integrate the color histogram as perceptual feature, shown in figure 2.

The energy of the image will be gathered into some DCT coefficients after DCT transformation, DC coefficients contain the main information of the original data matrix, AC coefficients contain the detail information of the data matrix. Meanwhile, they are the most sensitive information of human visual system.



$$PerFea = \begin{bmatrix} D^1_{1,gray}, D^1_{2,gray}, \cdots, D^1_{8,gray}, D^1_{1,fre}, D^1_{2,fre}, \cdots, D^1_{8,fre} \\ D^2_{1,gray}, D^2_{2,gray}, \cdots, D^2_{8,gray}, D^2_{1,fre}, D^2_{2,fre}, \cdots, D^2_{8,fre} \\ \cdots\cdots\cdots\cdots\ddots \\ D^{64}_{1,gray}, D^{64}_{2,gray}, \cdots, D^{64}_{8,gray}, D^{64}_{1,fre}, D^{64}_{2,fre}, \cdots, D^{64}_{8,fre} \end{bmatrix}$$

Figure 2. Process of perceptual feature extracting

2.3 PCA Based Inter-Quantization

Each column of perceptual feature (matrix) is an indicator, reflecting appropriate information of the input image. For example, the first column is constituted with each block's occurrences of pixels between 0 and 31, reflecting the distribution of the input image on the pixel interval. However, there is inherent correlation among adjacent blocks. So, each column of perceptual feature matrix contains some redundant information.

Principal Component Analysis is used to reduce, even elimilate, the redundant information in perceptual feature matrix. Then, the perceptual feature is compressed into a inter-feature matrix, with a smaller dimension (10*64) and few redundant information.

2.4 Hash Computation

Once the inter-feature matrix is generated, we can obtain the perceptual hash of input image via binarizing. Each column is binarized using the median of the rank-ordered coefficients.

IJCSI International Journal of Computer Science Issues, Vol. 10, Issue 1, No 3, January 2013
ISSN (Print): 1694-0784 | ISSN (Online): 1694-0814
www.IJCSI.org

If the subset of rank-ordered coefficients is denoted as $c(i), i \in 0,1,\cdots,K$, then their median is calculated as $\mu=(c(K/2)+c((K+1)/2))/2$. Then, the perceptual hash of input image is obtained by thresholding each column with the median μ as follows:

$$hash_i = \begin{cases} 1, & c(i) \geq \mu \\ 0, & c(i) < \mu \end{cases}, \; i \in 0,1,\cdots,K \qquad (2)$$

3. Experimental Results

The proposed algorithm has been implemented with matlab scripts. The evaluation was based on 72 distinct images. These images are all from corel image galley (URL: http://calphotos.berkeley.edu/).

The bit error rate(BER) [9] [10] is denoted as the rate of mismatched bits by comparing two perceptual hashes.

3.1 Robustness to Image Operation

Robustness implies that perceptual hash functions should be robust to all kinds of image operation (contrast increase, median filter, JPEG compression, noise addition, histogram equalisation, laplace sharpen, rotation), since the underlying content is never changed. That's to say, the BER between the perceptual hash of the raw image and the perceptual hash of the operated image should be infinitely close to 0.The experimental results of robustness evaluation is presented in table 1.

TABLE I. RESULTS OF ROBUSTNESS EVALUATION

Image Operation		***Mean of BER***
Contrast Increase	-30%	8.50%
	-20%	3.27%
	20%	4.10%
	30%	9.20%
Median filter		10.06%
JPEG Compression	10%	3.00%
	20%	8.22%
	40%	17.50%
Noise Addition	Gaussian	7.10%
	Peper and Salt	8.50%
Histogram Equalisation		11.27%
Laplace Sharpen		20.24%
Rotation	-5	23.39%
	-3	15.70%
	+3	17.33%
	+5	22.00%

Table 1 shows the mean of BER between the raw image's perceptual hash and the operated image's perceptual hash. Notice that most of the results are lower than 0.1 and stay close to the theoretical value 0, only when the operation (JGEP compression 40%, rotation 5) has changed the underlying content. Meanwhile, the more the underlying content changes, the bigger the mean of BER is. For example, the result is 0.0300 while the JGEP compression is 10%, and 0.1750 while the JGEP compression is 40%.

3.2 Discrimination to Different Images

Discrimination means that the hash functions are statistically independent for different perceptual content, so that any two distinct images result in different and apparently random perceptual hash.

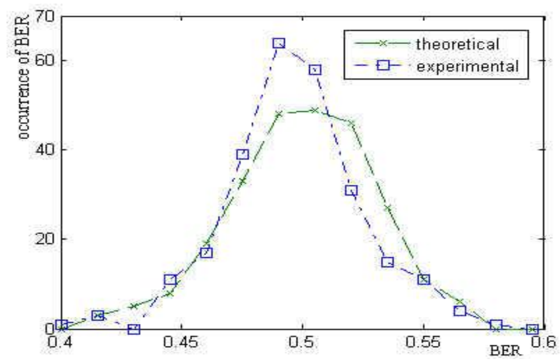


Figure 3. Distribution of BER between distinct images

Assume that two distinct images (i and i') are perceptually different, the theoretical optimal value of their BER $M_{BER}(pHash)$ can be estimated as follows:

$$M_{BER}(pHash) = E[BER(i,i')] \qquad (3)$$

where i and i' are taken independently and randomly from a given image set, and $E[]$ denotes mathematical expectation. According to Baris Coskun's and Kevin Hamon's analysis and proof in article [9] and [10], the theoretical optimal value is speculated to be 0.5.

Without loss of generality, we calculate BER between 72 test images, using 255 BER computation overall. Then we count the occurrence of each BER value and obtain the experimental distribution of BER, as shown in Figure 3. Meanwhile, we also plot the theoretical probability density function with $\mu = 0.5$ and $\sigma^2 = 0.0009$ in Figure 3.

The BER between perceptual hash of distinct images has a Gaussian distribution around the mean value of 0.4996, which is close to the theoretical optimal value 0.5. Thus, the perceptual hash of different images can be regarded as statistically independent as expectation.

3.3 Detection to the Tamper—Addition of a Logo

Tamper always brings in malicious changes to the original content of raw image. Typical image tampering operations include adding LOGO, image mosaic and so on. In this article, we mainly concern the tamper of adding LOGO, since the addition of a LOGO causes minimal change and is widely used with the increasing spread of Internet. The sample of adding LOGO is shown in Figure 4:

Figure 4. Sample of adding a LOGO

The content of tampered image shown in Figure 4 is much similar with the law image, only with malicious changes via adding a logo on the left. In order to detect such tampering operation to image, the BER between raw image and tampered image should be bigger than the BER of robust operation. Meanwhile, in order to distinguish tampered image from distinct image, the BER between raw image and tampered image should also be smaller than the BER of distinct images. We calculate 12 BER between raw image and tampered image, as shown in Figure 5:

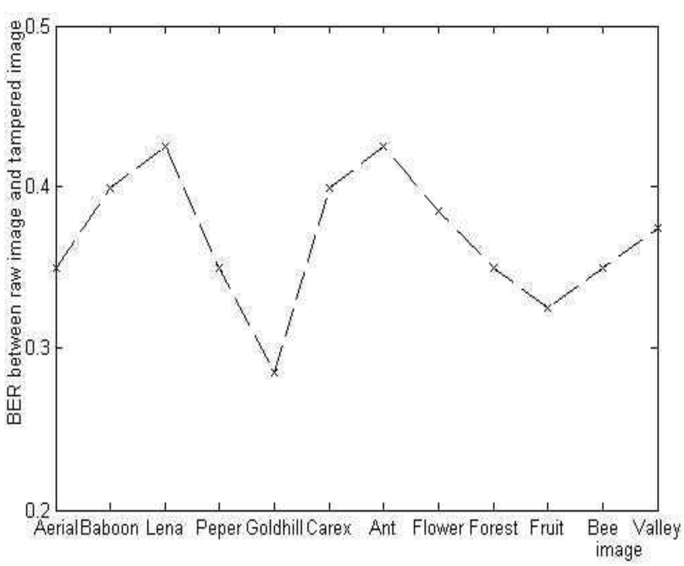


Figure 5. BER between raw image and tampered image

In Figure 5, most of the BER between raw image and tampered image is in the range of [0.3, 0.45]. Note that the most BER of robust operation is lower than 0.1 and the BER of distinct images has a Gaussian distribution around the mean value of 0.4996. Thus, the proposed image perceptual hash algorithm can effectively address such tamper detection problem with advantageous robustness and discrimination.

4 Conclusions

In this article we proposed a robust and discriminative image perceptual hash algorithm in order to address the problem of the tamper detection problem of images. We integrate color histogram and low-frequency DCT coefficients of image blocks as perceptual feature, then compress perceptual feature as inter-feature with PCA, and threshold the inter-feature to create a robust hash. Experimental results show that the proposed algorithm is advantageous at robustness since the most BER of robust operation is lower than 0.1, but it is a pity that the robustness toward rotation is not so perfect as assumed. It’s also found that the proposed algorithm has a very discriminative power because the BER of distinct images has a Gaussian distribution around the mean value of 0.4996, which is close to the theoretical optimal value 0.5. With such advantageous robustness and discrimination, the proposed algorithm can effectively detect the tampering operation of adding a LOGO.

Future investigation will address the problem of verification of the tamper detection ability toward other tampering operations. Meanwhile, we will extend this still-image perceptual hash method to video clip to address the problem of video authentication and copyright protection.

Acknowledgments

This work was partially funded by Natural Science Foundation of China through the Free Application Program under contract 61103174, Science and Technology Program of Shenzhen (China) through the Basic Research Program under contract JC201105170647A, and Laboratory and Device Management Research Foundation of Shenzhen University through the Basic Research Program under contract 2011045.

The Authors would thanks to Natural Science Foundation of China and Shenzhen Municipal Science and Technology Trade and Industry and Information Technology Commission (China) for their funding of our projects. The Authors also express gratitude to Laboratory and Device Management Department of Shenzhen University for their support and collaboration.

References

[1] NIU Xia-mu ,JIAO Yu-hua, "An Overview of Perceptual Ha shing," ACTA ELECTRONICA SINICA.China, Beijing, Vol.36 , No. 7, 2008, pp. 1405-1411.

[2] Bian Yang , Fan Gu , Xiamu Niu, "Block Mean Value Based Image Perceptual Hashing," International Conference on Intelligent Information Hiding and Multimedia Signal Processing(IIH-MSP '06), 2006, pp.167-172.

[3] J. Fridrich, "Visual hash for oblivious watermarking," Proc .IS&T/SPIE 12th Annu. Symp., Electronic Imaging, Security and Watermarking of Multimedia Content II, San Jose, CA, Jan. 2000, pp.286-294.

[4] R. Venkatesan, S.Koon, M. Jakubowski, and P. Moulin, "Robust image hashing," Proc. IEEE Int. Conf. Image Processing 2000, vol. 3, pp.664-666.

[5] M. K. MIhCak,R Venkatesan, "New Iterative Geometric Methods for Robust Perceptual Image Hashing," Proc of ACM Workshop on Security and Privacy in Digital Rights Management . Philadelphia :LNCS ,2001,pp.13-21.

[6] F.Lefbvre,B.Macq,J.-D.Legat, "RASH: RAdon Soft Hash algorithm," Proc. EUSIPCO,Toulouse,France,2002. pp.54-61.

[7] J.S.Seo, J.Haitsma, T.Kalker,C.D.Yoo,"A robust image fingerprinting system using the radon transform," Signal Process.: Image Commun., vol. 19, no. 4, 2004,pp. 325–339.

[8] Hui Zhang,Haibin Zhang,Qiong Li, Xiamu Niu, "A Multi-Channel Combination Method of Image Perceptual Hashing," Fourth International Conference on Networked Computing and Advanced Information Management,2008. NCM '08.Volume 2, 2008,pp.87-90.

[9] Baris Coskun, Bulent Sankur, Nasir Memon, "Spatio-Temporal Transform Based Video Hashing," IEEE Transactions on Multimedia,VOL. 8,NO. 6, 2006. pp. 1190-1208.

[10] Kevin Hamon, Martin Schmuker, Xuebing Zhou, "Histogram-based perceptual hashing for minimally changing video sequence," The Second IEEE International Conference on Automated Production of Cross Media Content for Multi-Channel Distribution , 2006.

Zeng Jie received his M. Eng. Degree in Signal and Information Processing from the Tianjin University, Tianjin, China in 2001. From 2001 to 2006, He worked as a software engineer in Huawei Technology Ltd. Now he is a lecturer in the College of Information Engineering, Shenzhen University, China. His current research intersts include wireless networks, wireless commmunication, and cooperative wireless networks.