EVALUATING ROBUSTNESS OF PERCEPTUAL IMAGE HASH FUNCTIONS

[1]N.Mounika,
[1]M.Tech, CNIS
[1]MVGR College of Engineering, India.

Abstract: ***Perceptual hashing focuses on the authentication and integration of multimedia data. Perceptual hash functions generate hash values for an image based on the human visualization. This function calculates same hash values of the same images and various hash values for several images. Then by using similarity function or distance the hash values will be compared, then it will decide if both the images are perceptually distinct or not. This paper presents a benchmarking for 4 various pHash functions. The Discrete cosine transformation (DCT) form hash, marr-hildreth operator form hash, radial variance form hash and block mean value form hash functions. The actual intention of this paper is to evaluate the robustness of four different pHash functions. Inorder to evaluate the robustness of pHash function for an image based on different image operations like resizing, JPEG compression, rotation or horizontal flipping is used.***

Index Terms - Perceptual hash function, discrete cosine transform, Marr-hildreth operator, Radial variance based hash, Block mean value based hash, Authentication

I. INTRODUCTION

Due to the promotion of digitalization as well as social media, the volume of multimedia information browsed has grown exponentially. Perceptual image hashing is technique used for authentication of multimedia data inspired from the cryptographic hash function. The encryption field is important in information security for enabling end to end security.But for the encryption of multimedia data the tradition cryptosystems are disadvantageous due to two reasons. The size of multimedia data is very large. Therefore the traditional cryptosystem directly takes maximum instance for the encryption of multimedia data. The other issue is that after decryption the decrypted image data may contain small distortion that is acceptable due to the characteristics of human perception. But in traditional cryptosystems the original image data must be exactly equal to the decrypted image data. Strong perceptual image hashing techniques has been recently developed as primary solution to overcome the above issues and have constituted the center of a testing creating research region to the academy and also the multimedia industry. Perceptual Image hashing capacities remove certain highlights from picture and a hash value is calculated depending on these features. Such functions have been developed to set up the "perceptual equality" of picture content. Image authentication is executed by looking at the hash estimations of the first image and the image to be validated. PHashes are required for having the capacity for getting on content-preserving manipulations and reject unauthorized modifications. The content-preserving manipulations include transmission errors, noise addition, compression and quantization, scaling, cropping, rotation and resolution reduction.

And the content changing manipulations include adding new objects, removing objects, shadow manipulations such as changing of light conditions, changing of image characteristics such as color, textures, and structures.

The paper aims to analyze the robustness of pHash functions. Let us consider functions: a discrete cosine transform (DCT) based, a marr-hildreth operator based, a radial variance based and block mean value based hash functions. We conducted experiments, inorder to evaluate the robustness of perceptual hash functions. To compare robustness of hash functions, we used speed, resizing, horizontal flipping, rotation or JPEG compression.

II. RELATED WORK

The researchers have proposed a variety of image hashing algorithms that extract features from an image. In [1] evaluates the robustness of different pHash algorithms based on the visual modification and wants a robust hashing algorithm to identify and track the images when they are uploaded in the social media like facebook, instagram or google [2] a watson's visual model is used for the extraction of sensitive features from an image then a robust pHash code is generated by integrating image block based features and key point based features. In perceptual image hash functions [3] compares four normalized block mean value derived image perceptual hashing functional algorithms and enhances the robustness of algorithms to geometric distortion. In [4] the author develops an algorithm which generates image hash by using fourier transform characteristics as well as controlled randomization and evaluated performance under several image processing functional operations and introduced a framework for studying and evaluating the security of image hashing systems and the framework model the hash values as random variables and quantify uncertainty in terms of differential entropy [5] uses secret key for randomly modulating image pixels to create the transformed feature space and then feature space is used for calculating the image hash and a 4-bit quantization scheme is used in the reduction of hash and this scheme reveals the robustness against JPEG compression, low pass and high pass filtering [6] proposes an iterative feature detector which extracts significant feature points in order to reduce vulnerability to adversarial attack probabilistic quantization is applied the algorithm withstands standard attacks including rotation, compression, scaling and signal processing operations [7] ring based entropies was used in the generation of hash the input images are converted into normalized images and then the normalized images are divided into rings then produces the hash from the ring based entropies correlation coefficient is used in the measurement of hash similarity [8] compares radial variance derived as well as marr-hildreth operator derived hash functions on biometric templates and evaluates the resistance to various types of attacks on biometric templates. The approach proposed in [9] combines color vector angles along with distinct wavelet transform the input image is modified to the normal size through Bicubic interpolation and blurred through gaussian filter then color vector angles are computed and partitioned into non overlapped a feature matrix is extracted and compressed by DWT which forms the image hash robust against JPEG compression, watermarking embedding and rotation within 5^0. Recent works focusing on the deep learning framework to produce binary hash codes for speedy image retrieval [10][11].

III. Perceptual image hash functions

A. DCT based hash

The DCT is one of the perceptual hash function used in generation of hash value from an image. The DCT is related to fourier transform, specifies a function or signal as sum of sinusoids with several amplitudes and frequencies. The DCT exercises cosine function whereas DFT uses both sine and cosine functions. The DCT hash eight various standard variations.

The type-II DCT is defined as,

$$\mathrm{Y[n]} = \sqrt{\frac{2}{\mathrm{N}}}.\sum_{\mathrm{m}=0}^{\mathrm{N}-1} \mathrm{y(m)}.\cos\left(\frac{(2\mathrm{m}+1).\mathrm{n}\pi}{2\mathrm{N}}\right) \qquad (3.1)$$

Where, n = 0…, N-1 and m= 0…., N-1

Also it is defined as following expression,

$$\mathrm{Y[n]} = \sum_{\mathrm{m}=0}^{\mathrm{N}-1} \mathrm{c[n,m].y[m]} \qquad (3.2)$$

Where c[n, m] represent the DCT matrix with ‘n’ rows and ‘m’ columns

The discrete cosine matrix c [n, m] is given as,

$$\mathrm{c\,[n,m]} = \sqrt{\frac{2}{\mathrm{N}}}.\cos\left(\frac{(2\mathrm{m}+1).\mathrm{n}\pi}{2\mathrm{N}}\right) \qquad (3.3)$$

where m, n=0......, N-1

The DCT utilizes various properties inorder to create perceptual hash function.

B. Marr-Hildreth Operator Based Hash

The marr-hildreth derived hash operates mainly based on the edge detection process. This function calculates hash value based on the information of the edges. The interested features could be color or texture yet basically luminance is used. It is also designated as Laplacian of Gaussian (LOG), which belongs to distinct laplacian filter.

$$\nabla^2 \mathrm{f_c(n,m)} = \nabla.\nabla \mathrm{f_c(n,m)} = \frac{\partial^2 \mathrm{f_c(n,m)}}{\partial \mathrm{n}^2} + \frac{\partial^2 \mathrm{f_c(n,m)}}{\partial \mathrm{m}^2} \qquad (3.4)$$

Where $\mathrm{f_c(n,m)}$ considered as an image’s grey level function.

Because of second derivative the edge points of $f_c(n,m)$ appear at the zero crossing of $\nabla^2 f_c$(n, m). The edge detection based on laplacian generates edges of null thickness. By using continuous laplacian different discrete laplacian operators are constructed. By using convolution a filter h $(x_{1,}\, x_2)$ applied for image space.

$$\nabla^2 f(x_1,x_2) = f(x_1,x_2) * h(x_1,x_2) \qquad (3.5)$$

In order to obtain edge map of an image, there requires one more process to be conducted. For a discrete space image $\nabla^2 f(x_1,x_2)$ the zero crossing has to be located. For this each pixel p in the image has to be compared with its eight neighbor’s pixels. If there is any sign difference in between the pixel p and the neighbor pixel q, then the edge lies between the p and the q. It is categorized as a zero crossing pixel if

$$|\nabla^2 f(p) \leq \nabla^2 f(q)| \qquad (3.6)$$

The Marr-Hildreth operator is as well known to be Laplacian of Gaussian (LOG). By applying the laplacian onto the gaussian the filter kernel is obtained. The gaussian filter is given as

$$g_c(n,m) = e^{-\frac{n^2+m^2}{2\sigma^2}} \qquad (3.7)$$

By interchanging laplacian and convolution,

$$\nabla^2[\mathrm{f_c(n,m)} * \mathrm{g_c(n,m)}] = [\nabla^2 g_c(n,m)] * f_c(n,m) \qquad (3.8)$$

The LOG filter $h_c(n,m)$ is given as,

$$h_c(n,m) = \nabla^2 g_c(n,m) = \frac{n^2+m^2-2\sigma^2}{\sigma^4}\;.\,e^{-\frac{n^2+m^2}{2\sigma^2}} \qquad (3.9)$$

C. Radial Variance Based Hash

The radial variance based hash function works based on the Randon transformation. The Radon transform is the integral transform over a straight line contains the integral of a function. Along the line projections it utilizes the difference rather than total of the pixel esteems. The variance catches luminance discontinuities along the projection lines much better. These discontinuities produce from edges that are perpendicular to the projection heading. The radial variance based hash is robust against various image processing stages such as compression and rotation. To discrete images in order to extend randon transform along the line integral $\mathrm{d = x.\cos\alpha + y.\sin\alpha}$ can be estimated by sum of the pixels.

$$\mathrm{d} - \frac{1}{2} \leq \mathrm{x.\cos\alpha + y.\sin\alpha} \leq \frac{1}{2} \qquad (3.10)$$

Where α is the angle of projection line used and x is the pixel coordinate at x-axis and y is the pixel coordinate at y-axis.

On projection line corresponding to the angle α, then $\tau(\alpha)$ defines the set of pixels (x, y). $(x,y) \in \mathrm{T}(\alpha)\;\; if$

$$-\frac{1}{2} \leq (x-x').cos\alpha + (y-y').sin\alpha \leq \frac{1}{2} \qquad (3.11)$$

Where (x',y') denote central pixels coordinates of an image.

The radial variance vector $V[\alpha]$ is given by,

$$V[\alpha] = \frac{\sum_{(x,y)\in\tau(\alpha)} I^2(x,y)}{\#\tau(\alpha)} - \left(\frac{\sum_{(x,y)\in\tau(\alpha)} I^2(x,y)}{\#\tau(\alpha)}\right)^2 \qquad (3.12)$$

Where α=0, 1, 2... 179 and I(x, y) designate the pixel (x, y) luminance value.

D. Block Mean Value Based Hash

The block mean value derived hash is one of the perceptual hash function used in the generation of the hash value from an image. In block mean value derived hash slightly 4 distinct functions were initiated. The first two functions depend on the normalized block mean values as well as other two functions are to increase the performance over rotation attacks.

Method 1 is Based on blocks mean value, the first method is described.

- The actual image is normalized into a predetermined sizes;
- Then the image G is divided into non overlapped blocks G1, G2, …, GN
- The block sequence indices {G1, G2, …,GN} are encrypted to {G’1, G’2, …, G’N }

- The mean value sequence {M1, M2, …, MN} is calculated from {G'1, G'2, …, G'N }

$$M_d = \text{median}(M_g) \quad (g=1,2,\ldots N) \quad (3.13)$$

- The mean value series is normalized in to binary method as well as a hash value h is obtained.

$$h(g) = \begin{cases} 0 & M_g < M_d \\ 1 & M_g \geq M_d \end{cases} \quad (3.14)$$

Method 2

- To calculate hash values method 1 is considered.
- The scale of overlapping is set to be the half-size of a block

Method 3-In this method, mean values of pixel blocks are rotated.

- The actual image is normalized into predetermined sizes;
- Then the image G is divided into non overlapped blocks G1, G2, …, GN
- The block sequence indices {G1, G2, …,GN} are encrypted to {G'1, G'2, …, G'N }
- The mean value sequence {M1, M2, …, MN} is calculated from {G'1, G'2, …, G'N }

$$M_d = \text{median}(M_g) \quad (g=1,2,\ldots N) \quad (3.15)$$

- Then matrix M established by {M1, M2, …, MN} rotated by D degrees, D = {0, 15, 30, … , 345}
- The rotated matrix Mg (g = 1, 2… 24) is divided in to N blocks.
- Then the mean value series {Mg1, Mg2... MgN} of each and every block and median value Mdg *o*f this series is obtained which creates 24 groups of series.
- The 24 groups of sequences are normalized into binary form and the ultimate hash value matrix is obtained.

Method 4 is combination of method 2 and method 3.

- As in the method 3 the mean values of pixels blocks are rotated and accomplish overlapping operation same as in Method 2.
- Then the hash value is computed from its rotated values of the blocks.

IV. RESULTS AND DISCUSSION

4.1 Speed:

The perceptual image derived hash function is crucial during which a amount of images need to be hashed as well as processed. In order for evaluating the speed of the perceptual hash functions, let us consider an image set "event image set" consisting of 47 images with distinct modifications. The mean dimension of the images is 2874x2260 pixels and the file size is 3.25MiB. And the size of file containing total 47 images is 150.80MiB.Hence each pHash has to hash 47x2=94 images. First of all to the attack chain an "empty" type of attack is added. Then the images present in image directory were copied into attack image directory without changing them. In the four different perceptual hash functions the newly developed block mean value based hash needs 60 seconds to hash total 94 images which is the fastest hash function. The results are summarized as,

TABLE 4.1-Speed metrics

	DCT	MH	Radial variance	BMH
Total time in seconds	922	351	121	60
Average seconds per image	8.9	3.4	1.6	0.9
MiB per second	0.45	0	2.45	5.43

And the radial variance based hash which requires 121 secs to hash 96 images is the 2^{nd} fastest algorithm. The Marr-hildreth (351) and the DCT (922) are the far behind hash functions. The variance in speed results due to the block mean value based hash and radial variance hash functions particularly utilize pixel operation for feature extraction while hash is calculated. The remaining marr-hildreth and DCT use computationally expensive convolution/correlation operations.

4.2 Intra Score Distribution:

In order to try out the robustness of the perceptual image hash derived functions various common multimedia operations were performed. The common image operations include JPEG compression and rotation for decreasing the file size of an image and to modify their images. In order to demonstrate the robustness in scientific papers rotation operation is used. The human perception of an image is hardly changed due to horizontal flipping so it is also considered.

To evaluate score distribution, consider a "chaos image set" consisting of 45 images for a mean dimension of 2502x2200 pixels. The mean file size is 2.5MiB and the total file size of 45 images is 110.10MiB. The set of images which is the subset of chaos image set is called "small chaos image set" is also taken. To form this image set 3 images were considered from the chaos image set and the mean dimension of images is 3003x2222 pixels. The mean file size is 3.51MiB and the complete size of 3 images is 11.54MiB. The used image sets are converted from JPEG format to PNG format in order to overcome the lossy compression. Therefore the chaos image set complete file size increases from 109.90MiB to 343.85MiB and the mean file size increases from 2.5to 7.70MiB. And the size of the small chaos image set increases from 11.54MiB to 35.90MiB. Therefore chaos image set is utilized for score distribution and the small chaos image set is for attack charts.

4.3 Horizontal Flipping:

The binary representation of an image changes drastically, if an image is flipped. Through the perception of the human visual analysis the meaning of the image changes minimally or not at all. The results are,

TABLE 4.2-Horizintal Flipping

	DCT	MH	Radial	BMB
Mean distance	0.479	0.483	0.499	0.315
Maximum distance	0.625	0.658	0.732	0.703
Minimum distance	0.375	0.276	0.042	0.047

By horizontal flipping the images were changed. In the four different perceptual hash functions none of the function is robust against horizontal flipping.

4.4 Resizing:

Resize the width to 1024 pixels such that the images were changed and proportionally the height is arranged. Bicubic interpolation was used. In the four different perceptual hash functions the radial variance derived hash function is not robust to resizing. It is due to the fact that before extracting its features it does not normalize the resolution of an image.

TABLE 4.3-Resizing

	DCT	MH	Radial	BMB
Mean distance	0.076	0.068	0.348	0.012
Maximum distance	0.219	0.271	0.670	0.039
Minimum distance	0.000	0.017	0.008	0.000

4.5 JPEG Compression:

The robustness of perceptual hash functions can be obtained by JPEG quality setting. By changing JPEG quality setting to 80 the images was changed. Therefore, by varying the JPEG quality setting from 100 to 0. For every value the hash function should calculate distance score for the images.

TABLE 4.4-JPEG compression

	DCT	MH	Radial	BMB
Mean distance	0.003	0.055	0.000	0.002
Maximum distance	0.021	0.345	0.000	0.007
Minimum distance	0.000	0.003	0.000	0.000

For JPEG compression, the radial variance based hash is more robust. Even for the JPEG quality setting of 0, the distance score is negligible for radial variance based hash. Up to quality setting of 10 the DCT and block mean value based hash also performs same as radial variance based hash. The marr-hildreth derived hash performs most awful in this case.

4.6 Rotation:

Inorder to evaluate the robustness of hash functions regarding rotation the Bicubic interpolation is used and the images were rotated by 5^0. By varying the angle of rotation gradually the results were summarized. While performing rotation operation on images, none of the hash function performs efficiently. But when the images are rotated by 3^0 then the block mean value derived hash performs best. The marr-hildreth derived hash won't perform efficiently even if the image is rotated by 1^0.

TABLE 4.5-Rotation

	DCT	MH	Radial	BMB
Mean distance	0.326	0.466	0.346	0.230
Maximum distance	0.601	0.565	0.781	0.364
Minimum distance	0.071	0.348	0.052	0.118

V. CONCLUSION

This paper evaluates the robustness property of following pHash functions: DCT derived hash, marr-hildreth operator form, radial variance form and block mean value derived hash. It evaluates robustness of perceptual hash functions based on different image operations like JPEG compression, speed, rotation, resizing and horizontal flipping. To evaluate robustness different image sets were which are obtained from Wikimedia commons. To compare robustness of different hash functions we used common measures: Mean distance, Maximum distance and Minimum distance.

The evaluation results obtained confirm that block mean value derived perceptual hashing is the quickest and robust against JPEG compression, rotation and resizing compared with other perceptual hash functions. Regarding horizontal flipping no other tested hash function is robust.

REFERENCES

[1] Drmic, Andrea, et al. "Evaluating robustness of perceptual image hashing algorithms." *Information and Communication Technology, Electronics and Microelectronics (MIPRO), 2017 40th International Convention on*. IEEE, 2017.

[2] Wang, Xiaofeng, et al. "A visual model-based perceptual image hash for content authentication." *IEEE Transactions on Information Forensics and Security* 10.7 (2015): 1336-1349.

[3] Yang, Bian, Fan Gu, and Xiamu Niu. "Block mean value based image perceptual hashing." Intelligent Information Hiding and Multimedia Signal Processing, 2006. IIH-MSP'06. International Conference on. IEEE, 2006.

[4] Swaminathan, Ashwin, Yinian Mao, and Min Wu. "Robust and secure image hashing." *IEEE Transactions on Information Forensics and security* 1.2 (2006): 215-230.

[5] Ahmed, Fawad, Mohammed Yakoob Siyal, and Vali Uddin Abbas. "A secure and robust hash-based scheme for image authentication." *Signal Processing* 90.5 (2010): 1456-1470.

[6] Monga, Vishal, and Brian L. Evans. "Perceptual image hashing via feature points: performance evaluation and tradeoffs." *IEEE Transactions on Image Processing* 15.11 (2006): 3452-3465.

[7] Tang, Zhenjun, et al. "Robust image hashing using ring-based entropies." *Signal Processing* 93.7 (2013): 2061-2069.

[8] Vukelic, Bernard, and Miroslav Baca. "Comparison of RADIAL variance based and Maar-Hildreth operator perceptual image hash functions on biometric templates." *Central European Conference on Information and Intelligent Systems*. Faculty of Organization and Informatics Varazdin, 2014.

[9] Tang, Zhenjun, et al. "Robust image hashing via colour vector angles and discrete wavelet transform." *IET Image Processing*8.3 (2013): 142-149.

[10] Lin, Kevin, et al. "Deep learning of binary hash codes for fast image retrieval." *Computer Vision and Pattern Recognition Workshops (CVPRW), 2015 IEEE Conference on*. IEEE, 2015.

[11] Xia, Rongkai, et al. "Supervised hashing for image retrieval via image representation learning." *AAAI*. Vol. 1. 2014.

[12] Zauner, Christoph. "Implementation and benchmarking of perceptual image hash functions." (2010).

[13] Hadmi, Azhar, et al. "Perceptual image hashing." *Watermarking-Volume 2*. InTech, 2012.

[14]Kessler, Gary C., and Chet Hosmer."An overview of steganography." *Advances in Computers*. Vol. 83. Elsevier, 2011. 51-107.

[15] Lin, Ching-Yung, and Shih-Fu Chang. "A robust image authentication method distinguishing JPEG compression from malicious manipulation." *IEEE Transactions on Circuits and Systems for Video Technology* 11.2 (2001): 153-168.

[16] Kozat, Suleyman Serdar, Ramarathnam Venkatesan, and Mehmet Kivanç Mihçak. "Robust perceptual image hashing via matrix invariants." *Image Processing, 2004. ICIP'04. 2004 International Conference on*. Vol. 5. IEEE, 2004.

[17] Han, Shui-Hua, and Chao-Hsien Chu. "Content-based image authentication: current status, issues, and challenges." *International Journal of Information Security* 9.1 (2010): 19-32.

[18] Mıhçak, M. Kıvanç, and Ramarathnam Venkatesan. "New iterative geometric methods for robust perceptual image hashing." *ACM Workshop on Digital Rights Management*. Springer, Berlin, Heidelberg, 2001.

[19] Shuo-zhong, Wang, and Zhang Xin-peng. "Recent development of perceptual image hashing." *Journal of Shanghai University (English Edition)* 11.4 (2007): 323-331.

[20] Monga, Vishal. *Perceptually based methods for robust image hashing*. Diss. 2005.

[21] Tang, Zhenjun, et al. "Perceptual image hashing with histogram of color vector angles." *International Conference on Active Media Technology*. Springer, Berlin, Heidelberg, 2012.

[22] Xiang, Shijun. "Histogram-Based Perceptual Image Hashing in the DWT Domain." Multimedia Information Networking and Security (MINES), 2010 International Conference on. IEEE, 2010.

[23] Khelifi, Fouad, and Jianmin Jiang. "Perceptual image hashing based on virtual watermark detection." *IEEE Transactions on Image Processing* 19.4 (2010): 981-994.

[24] Sebastian, Lima S., Abraham Varghese, and T. Manesh. "Image authentication by content preserving robust image hashing using local and global features." *Procedia Computer Science* 46 (2015): 1554-1560.

[25] Qin, Chuan, et al. "A novel image hashing scheme with perceptual robustness using block truncation coding." *Information Sciences* 361 (2016): 84-99.

[26] Jothimani, S., and P. Betty. "A survey on image authentication techniques." *Int. J. Eng. Trends Technol.* 7.4 (2014): 184-186.

[27] Mishra, Manish, and Shraddha Pandit. "IMAGE ENCRYPTION TECHNIQUE INCORPORATING WAVELET TRANSFORM AND HASH INTEGRITY."

[28] Qin, Chuan, Shuo-zhong Wang, and Xin-peng Zhang. "Image hashing based on human visual system [J]." Journal of Image and Graphics 11.11 (2006): 1678-1681.

[29] Lv, Xudong, and Z. Jane Wang. "Perceptual image hashing based on shape contexts and local feature points." *IEEE Transactions on Information Forensics and Security* 7.3 (2012): 1081-1093.

[30] Lei, Yanqiang, Yuangen Wang, and Jiwu Huang. "Robust image hash in Radon transform domain for authentication." *Signal Processing: Image Communication* 26.6 (2011): 280-288.

[31] Ouyang, Junlin, Gouenou Coatrieux, and Huazhong Shu. "Robust hashing for image authentication using quaternion discrete Fourier transform and log-polar transform." *Digital Signal Processing* 41 (2015): 98-109.

[32] Eskenazi, Sébastien, et al. "A perceptual image hashing algorithm for hybrid document security." *Document Analysis and Recognition (ICDAR), 2017 14th IAPR International Conference on*. Vol. 1. IEEE, 2017.