International Journal of Engineering Research-Online
A Peer Reviewed International Journal
Articles available online http://www.ijoer.in

Vol.3., Issue.4., 2015
(July-Aug)

RESEARCH ARTICLE

ISSN: 2321-7758

A NOVEL IMAGE HASHING ALGORITHM BASED ON SALIENT REGION DETECTION

CHUNXING.WANG[1*], MIAOMIAO.TONG[2]
[1*]Professor, School of Physics & Electronics, ShangdongNormal University.
No.88 of East Wenhua Road, Jinan, 250014, P.R. China
Email: cxwang@sdnu.edu.cn
[2]Student, School of Physics & Electronics, ShangdongNormal University.
No.88 of East Wenhua Road, Jinan, 250014, P.R. China
Email: tongmiaomiao92@163.com

CHUNXING.WANG

ABSTRACT
Image hashing is the process of encoding digital image data into compact binary codes. It is useful for applications like image authentication, digital watermarking, and content-based image retrieval. A novel image hashing method is presented in this paper, which is developed using the detection of visually salient image regions. This method includes selecting a specific frequency band for the Gauss difference filtering, converting RGB channels of a host image into the intended channels, computing saliency and producing the hash sequence. Our experiment demonstrates the effectiveness of the proposed algorithm and shows that the method is robust to various types of attacks such as image zooming, rotating, Gauss blur, histogram equalization and image median filtering and so on. At the same time the algorithm is superior in image tailoring, image mean filtering and median filtering compared with the contrast algorithm.

Keywords—image hashing, salient region, color spaces, DoG filter.

INTRODUCTION

In recent years, the scale of network multimedia data is growing at an exponential speed. With the popularity of digital cameras and smart phones, images are the most exposed multimedia data in daily life. An important requirement of the application of digital image is to realize fast image retrieval and image authentication for the large scale data image. In various coping techniques, image sensing is the hotspot in the field of multimedia security and multimedia information processing due to its robustness and copyright protection[1].Image hashing algorithm is based on the image content extracted from the digital image data and mapped into a fixed number string.

In general, image hashing algorithm includes image preprocessing, feature extraction, quantization and image matching. In the past decade, a lot of implementations have been presented for the applications of different engineering purposes and different robustness. Swaminathan[2]proposed a perceptual hashing algorithm based on the rotation invariance of Fourier-Mellin transform. Although the algorithm has good robustness to geometric transformation, but it is not ideal for the effect of additive noise attack. Khelifi[3]proposed an algorithm of extracting the hash with a virtual optimal watermark detector. This method can control the distribution of the hash and improve the security effectively. Z. Tang[4], Monga V[5], Xiang Shi-jun[6], et al. constructed robust image hashing via non-negative matrix factorizations.

It is robust to common image operations (such as lossy compression, low-pass filtering, scaling and so on). SUN Rui[7]presented a perceptual image hashing method via scale invariant feature transform (SIFT) and principal component analysis (PCA).This method is robust to various types of attacks such as image rotation, illumination change and filtering, etc. But these methods company with high complexity of the algorithm. Zhen Liu[8]proposed an algorithm to represent feature's spatial context information with a binary code. It is a contextual hashing for large-scale image search. Bao-lin SONG[9] proposed an image hash authentic scheme based on composite domain. The obtained hash is stable to median filter, mean filter, Gaussian noise and JPEG coding, while less sensitive to the malicious tampering.

This paper summarizes the existing digital image hashing algorithms, and proposes a novel image hashing scheme based on salient region detection, which is robust to image zooming, rotating, Gauss blur, histogram equalization and image filtering. The lower complexity of our algorithm is easy to realize, and proves to be more time efficient in the experiment.

Image feature selection

Among the hash extraction scheme, the most critical step is the extraction of the image features, which forms the core of the image hashing algorithm. In this paper, the feature vector is extracted from the salient region to construct the image hash.

Image data information can be classified into visually salient regions and non-visually salient regions[10]. Human vision system can pay close attention to the important part of the image and distinguish the salient regions of the image. In order to improve the accuracy of the analysis and reduce the amount of data processing, the detection of the saliency regions of the image becomes an important way for large data filtering. Detection of visually salient image regions can be widely used in the field of computer vision, including image segmentation for region of interest, object recognition, adaptive compression, image scaling and other fields. This paper can improve the accuracy of the image analysis and reduce the amount of data processing by using the image saliency maps, and obtain the image hash sequence more quickly.

DoG(Difference of Gaussian) Filter

DoG is the differential of Gauss function, proposed by D.G. Lowe. The DoG operator has strong robustness to noise, scale affine change and rotation. We can get the low-pass filtering result of an image by using the convolution between the image and the Gauss function, that is, the process of image de-noising. However, the DoG operator suppresses the high frequency and the low frequency components by the Gaussian Filter, forming a Band-Pass Filter. The band-pass means that only the designated frequency band can be passed, and in the spatial domain, manifests as only the specified scale detail of image can be retained. This is better than the ordinary high-frequency sharpening algorithm in detection details, because the ordinary high-frequency sharpening not only enhances details, but also enhances high-frequency noise. In addition, The DoG filter efficiently approximates the Laplace of Gaussian (LoG) filter and is more convenient and faster. On a certain scale, the feature detection can be achieved through the subtraction of two adjacent dimensions of Gaussian smoothed image, while it is worked as DoG response value of image. DoG operator is widely used in feature detection, especially in the detection of intensity transform. DoG has usually also been used for interest point detection and saliency detection[11]. The image $I(x,y)$ is filtered by DoG Filter with different ratio which can be represented by $f\ (x,y,\rho)$ in a scale-space, written by the following formula:

$$\begin{aligned} f\ (x,y,\rho) &= \frac{1}{2\pi}\left[\frac{1}{\sigma_1^2}e^{-\frac{(x^2+y^2)}{2\sigma_1^2}} - \frac{1}{\sigma_2^2}e^{-\frac{(x^2+y^2)}{2\sigma_2^2}}\right] \\ &= \big(g(\sigma_1) - g(\sigma_2)\big) * I(x,y) \\ &= G(x,y,\sigma_1) - G(x,y,\sigma_2) \quad (1) \end{aligned}$$

Where σ_1 and σ_2 are standard Gaussian mean square deviations $(\sigma_1 > \sigma_2)$, ρ represents the multiple of two adjacent scale spaces $(\rho=\sigma_1:\sigma_2)$."*" denotes the convolution operation, and $g(\sigma_1)$ - $g(\sigma_2)$ realizes the DoG filter. The pass-band width is controlled by the ratio $\sigma_1:\sigma_2$. When ρ is 1.6, the DoG filter efficiently approximates the Laplace of Gaussian (LoG) filter; and when the ratio ρ is 5, it approximates well the visual field of retinal ganglion cells. In order to obtain high quality visually salient image regions, the high frequencies arising from texture, noise and blocking artifacts need to be removed[12].Then the smaller Gauss fuzzy radius is needed which can reduce the amount of calculation.

Color space

Color space (also called color model or color system) is the standard of coordinate system and subspace, the mathematical representation of a set of colors. In general, color space can be classified into three categories: human visual system (HVS) (e.g. RGB, HSI, etc.); application specific (e.g. CMY, YUV, YCbCr, etc.) and CIE color space (e.g. CIELab)[13]. RGB model is the most common hardware-oriented model, which is generally used for color monitor and color video camera. Red (R), green (G) and blue (B) are the three primary additive colors, individual components are added together to form a desired color. RGB color space uses three primary physical colors, as a result, this color space has definite physical meaning, which is suitable for color kinescope work. However, RGB is not very efficient when dealing with 'real-world' images[14] or human visual features. Therefore the other color space representing methods are produced. HIS(Hue, Saturation and Intensity) color space reflecting the color of human visual system uses three basic features including hue (H), saturation (S) and intensity (I) to perceive color. Hue and saturation are commonly called chroma, and they are used to represent categories and shades of color. When processing color images, one can only deal with intensity (I) component, and the results do not change the color types of the original images. In order to facilitate the color processing and recognition, HIS color space is used frequently for the human visual system which is more sensitive to the brightness than the degree of tint. HIS model is more close to the way of human description and interpretation of color than RGB model. CMY model is usually for color printers, which is complementary to RGB space, that is to say, a white minus a RGB space in a color value equals the same color value in CMY space. Because of the chemical characteristic of color ink and pigment, black obtained by CMY is not really black. A real black is often added in printing, so CMY is written as CMYK. CMYK forms the so-called "color printing" by the superposition of cyan, magenta, yellow and black ink. CIE color space creates a new color system by obtaining the theoretical trichroism derived from the true primary colors via mathematical method. CIE-LAB model applies opponent color coordinate, which uses 'L', 'a', and 'b' coordinates to define the CIE color space. 'L' represents the brightness of the light, which ranges from 0 (black)-100 (white), 'b' and 'a' represent the chromaticity coordinates, with 'a' representing the red green axis, and 'b' representing yellow blue axis, both ranging from 0 - 10. CIE-XYZ is the foundation of other color systems, which uses red, green and blue as the three primary colors, and all other colors are derived from these three colors. RGB model uses three primary physical colors; it is the color space of device-dependent, while CIE model is device-independent. RGB model cannot be converted to CIE-Lab system directly, it needs to be converted to CIE-XYZ model before converting to CIE-Lab system.

Computing saliency

We first pre-filter the target image by differential filtering technique through DoG Filter. Then convert RGB channels into the intended ones. Finding the saliency information *S(x, y)* for an image I of width M and height N pixels in our paper can be formulated as:

$$S(x,y) = \left\| I_f - I_m \right\| \qquad (2)$$

Where I_f is the image feature vector value of the defining color space under the Gaussian blurred version (using a 3×3 separable binomial kernel) of the original image, I_m is the mean of image feature vector, and $\| \quad \|$ stands for the Euclidean distance. If the color model is chose CIE Lab model as the intended channels, the formula (2) can be rewritten as:

$$S(x,y) = (l - l_m)^2 + (a - a_m)^2$$
$$+ (b - b_m)^2 \qquad (3)$$

Where "*l*" denotes the feature vector value of the axis of the brightness, "*a*" stands for the feature vector value of the axis of red-green and "*b*" expresses the feature vector value of the axis of yellow-blue respectively; "l_m", "b_m" and "a_m" stand for the mean of the feature vectors respectively.

Extraction of hash sequences

Based on the above description, a significant information map of the size of M×N is obtained. We present a novel strategy for the extraction of hash sequences based on the saliency map $S(x,y)$, which is described as following and we use Fig.1 and Fig.2 to facilitate the understanding of the algorithm:

(i) Calculate the average of $S(x,y)$ and record as $\bar{Z}$.

(ii) Divide $S(x, y)$ into four quadrant regions on average, so the image size of each quadrant is $M/2 \times N/2$; each subset is denoted as z_i, and i=0, 1, 2, 3.

(iii) Compute the mean of each subset and denote as $\overline{z_i}$, i=0, 1, 2, 3. Get the first set of hash sequences H_1 by (4) which is shown as following:

$$h_i = \begin{cases} 1, & if\ \overline{z_i} \geq \bar{z} \\ 0, & Otherwise \end{cases} \qquad (4)$$

The length of this sequence is 4.

(iv) If $\overline{z_i} \geq \bar{z}$, then segment the subset z_i into four quadrant regions on average just like step (ii). Each sub-subset can be expressed as $z_{(i,j)}$, i=0,1,2,3; j=0,1,2,3. Otherwise stop the segmentation.

(v) Figure out the mean of the sub-subsets $z_{(i,j)}$ and denote as $\overline{z_{(i,j)}}$. The new hash sequence H is obtained by (4) and replace $\bar{z}$ by $\overline{z_i}$. Place H behind 1 in the H_1, and put "0000" behind 0 in the H_1. At this point, we have a set of binary sequences of the length 20. This process is presented in Fig. 1 and Fig. 2.

(vi) Repeat steps (iv) and (v), we can get the hash sequence of length L. The length L can be calculated by the following formula:

$$L = \sum_{i=1}^{N} 4^i \qquad (5)$$

Where N indicates the times of segmentation.

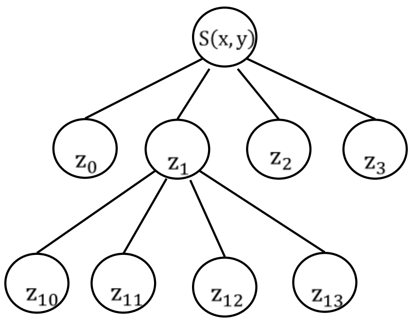


Figure 1 the quaternary tree of image segmentation

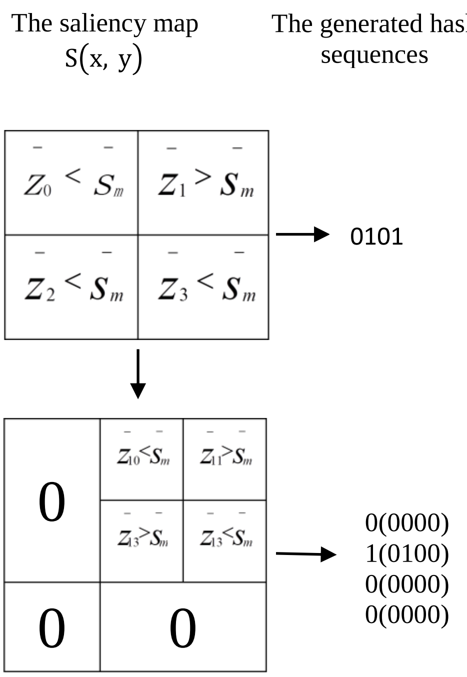


Figure 2 Schematic diagram of hash extraction algorithm based on S(x,y)

Experimental results

a) Selection of color space

The proposed image feature extraction scheme was implemented to evaluate the best color space for our hashing algorithm. Three test images: bird, house and girl are shown in Fig.3a-c.Fig.4 shows the saliency maps after CMYK color space, HIS, and CIE-Lab color space, respectively. We use the same size of the Gauss filter template and the same ratio ρ. From the results, it can be found that we can get high quality saliency maps based on the CMYK color, but not as good as based on the CIE-Lab. While there are obvious blocking artifacts in the HIS color space. By comparing all color spaces from the results, it can be found that the saliency information in CIE-Lab color space is greater than the others. It disregards high frequencies arising from texture, noise and blocking artifacts. So in the next experiment, we choose the CIE-Lab color space to obtain the saliency maps.


(a) Bird (b) House (c) Girl

Figure 3 test images

(a)CMYK (b) HIS (c) CIE Lab

Figure 3 saliency maps

b) Robustness evaluation

In this paper, 40 color images are selected as test images from the image database USI-SIPI of University of Southern California and the ground truth database from the University of Washington. We registered the hash value of each image in the database, and processed the original image in different ways and different degrees. Test the resistance of our algorithm for the attacks of common signal processing.

TABLE I. The generating method of attacked images

Attack types	intensity	step length
Gauss filter	[0.02-0.20]	0.02
Median filter	[3-21]	2
Histogram equalization	[2-256]	16
Scale transformation	[0.2-2.0]	0.2
Clipping	[0.05-0.30]	0.05
Rotation	[5-45]	5

The way given by Table 1 is undergone for the robust attacks, including image scaling, median filtering, Gauss blur, histogram equalization, rotating and clipping. According to the data listed in the table, each image is generated by a total of 53 copies of different forms.

Our paper mainly uses the hamming distance to measure the similarity between the hash sequences of origin images and the attacked images. The formula of hamming distance is given by (6).

$$D(H_1, H_2) = \frac{sum(H_1 \oplus H_2)}{dim(H_1)} \quad (6)$$

Where $\oplus$ indicates exclusive OR, sum( ) means summation, and dim( ) means dimension. According to the similarity value and a preset threshold to judge whether the image is the registered image. We obtained 40 hash sequences of the test images by the proposed hash extraction algorithm, calculated the hamming distance among the 40 hash sequences, and fix the threshold for the mean of the 780 hamming distances, which is 0.20 in this paper.

The simulation experiments are carried out based on Matlab,5%, 10% and 15% are worked as cutting ratios of the selected images, of the selected images are cut, calculated the average value of the hamming distance between the attacked images and the origin images, and compared with Ref. [9]. In the filtering operation, the selection of the common mean filter and median filter is tested. We test image filtering of the mask using 3×3 median filter and mean filter operation, respectively, calculate the average hamming distance and compare with Ref. [9].The results are shown in Table 2.

TABLE II. Performance comparison of contrast algorithm

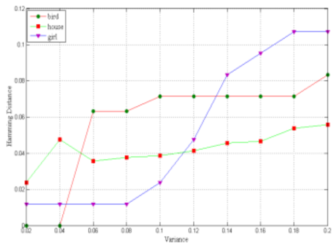
Attack types	Our algorithm	Algorithm in Ref. [9]
Clipping 5%	0.0286	0.1226
Clipping 10%	0.1726	0.1878
Clipping 15%	0.1931	0.2035
Mean filter [3 3]	0.0476	0.0301
Median filter [3 3]	0.0238	0.1021

From Table 2, we can find that the robustness of our algorithm is getting worse with the increase of the intensity of the image clipping, but it is better than the contrast algorithm. And the proposed algorithm has good robustness against image mean filtering and median filtering, which is much better than the contrast algorithm.

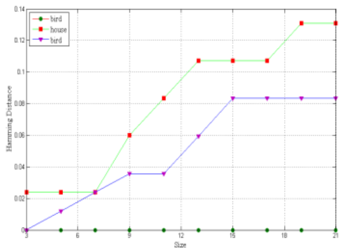
In order to test the robustness of the algorithm to conventional attacks, in our paper we take color images bird, house and girl as examples to illustrate the performance of the image hashing algorithm based on image saliency map.

The impact of the Gauss filter is shown in Fig.5 (a). We can see that even under the Gauss filter with the variance of 0.2, the hamming distance does not exceed the threshold, which shows that our algorithm has good robustness to Gauss noise. Fig.5 (b) shows the test results under the median filter. We can see from Fig.5 (b) that all the hamming distance is less than the threshold. Especially for bird image, this attack has no effect on it. So the proposed method has great robustness to the median filter. Fig.5 (c) shows that our algorithm is not sensitive to histogram equalization. But when the step is 256, the hamming distance of the image of house is 0.2432, this is because in the large extent, histogram equalization makes color image distortion. Fig.5 (d) illustrates the impact under scale transformation. It shows our algorithm has good robustness to image scale transformation.

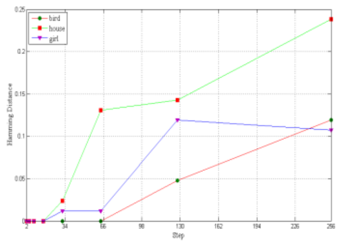
As shown in Fig.5 (e) and Fig.5 (f), the performance of our algorithm is poor for geometric transformation based on content. When the image is clipped 10%, the hamming distance of house image exceeds the threshold, and the hamming distance of girl image exceeds threshold when the image is clipped 15%. Fig.5 (f) illustrates that the image is very sensitive to the rotation operation. The house image can't bear the rotation operation and girl image can only resist small angle rotation. But we can see that bird image has good robustness to whether clipping or rotation. It benefits from that the salient region of bird image is very prominent. In the future, we will focus on the study of the optimized salient regions extraction algorithm.



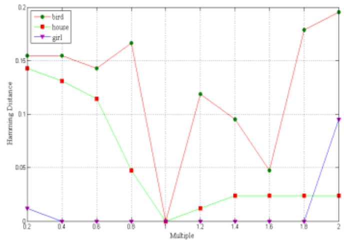
(a)



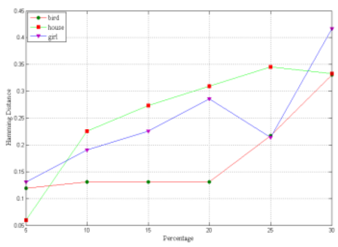
(b)



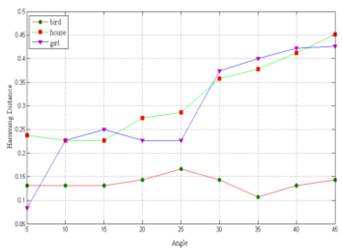
(c)



(d)



(e)



(f)

Figure 5 the impact of hamming distance under different attacks. (a) The performance under different variances of Gauss filters. (b) The performance under different size of the Median filters. (c) The performance under different step of the Histogram equalization. (d) The performance under different multiple of Scale transformation. (e) The performance under different percentage of Clipping. (f) The performance under different angle of Rotation.

Conclusion

In this paper, a new image hashing algorithm is designed by the detection of visually salient image regions. We use the image saliency map to improve the accuracy of analysis and reduce the amount of processing data. Simulation results prove that the algorithm has good robustness to Gauss filter, median filter, mean filter, histogram equalization and scale transformation. But our image hash is very sensitive to the change of image content, so the algorithm cannot resist the large scale image rotation attack and image clipping. In the future we may combine with other methods to resist the rotation attack and clipping. We will also apply the algorithm to image retrieval in various fields, such as book searching in the library.

REFERENCES

[1]. Wang S Z, Zhang X P.Recent development of perceptual image hashing[J].Journal of Shanghai University (English edition),2004,11(4):323-331.
[2]. Swaminathan A, Mao Y, Wu M. Robust perceptual image hashing via matrix invariants [A]. Proceedings of 2005 ICIP [C].Genova, Italy, 2005.3443-3446.
[3]. Khelifi F ,Jiang Jian min. Perceptual image hashing based on virtual watermark detection [J].IEEE Trans .on image processing,2010,19(4); 981-994.
[4]. Z.Tang,S.Wang,X.Zhang,W.Wei,S.Su,Robust image hashing for tamper detection using non-negative matrix factorization, Journal of Ubiquitous Convergence and Technology,vol.2,no.1,pp.18-26,2008.
[5]. Monga V, Mihcak K M.Robust and secure image hashing via non-negative matrix Factorizations [J].IEEE Trans. on information forensics and security, 2007, 2(3):376-390.
[6]. Xiang Shi-jun, Yang Jian-quan. NMF-Based Image Hashing Algorithm Using Restricted Random Blocking[J]. Journal of Electronics & Information Technology,2011,33(2):337-341.
[7]. SUN Rui, YAN Xiao-xing, Gao Jun. A perceptual image hashing method via SIFT and PCA, Journal of Circuits and Systems, Vol.18 No.1, 1007-0249 (2013) 01-0274-05.
[8]. Zhen Liu, Houqiang Li. Contextual Hashing for Large-Scale Image Search, IEEE Transactions on image processing, vol. 23, no. 4, pp.1606-1614, 2014
[9]. Bao-lin SONG, Chun-xing WANG. Image Hash Authentic Scheme Based on Composite Domain.2014 International Conference on Mechanical Engineering and Automation[ICMEA 2014].Jan,2014:268-273.
[10]. Li Chong fei, Gao Ying hui. Saliency detection method based on phase spectrum and amplitudespectrum tuning, Journal of Image and Graphics. 2012,Vol.17,No.7.
[11]. J. Harel, C. Koch, and P. Perona. Graph-based visual saliency. Advances in Neural Information Processing Systems,19:545–552, 2007.
[12]. RadhakrishnaAchantay. Frequency-tuned Salient Region Detection, CVPR.Page 4-5, 2009.
[13]. Mehdi Khalili, David Asatryan. Colour spaces effects on improved discrete wavelettransform-based digital image watermarking using Arnold transform map. IET Signal Process,2013,Vol.7,Iss.3,pp.177-187.
[14]. Kim, S.H. Lee, H. Lee, Y.H. Ha H.Y.: 'Digital watermarking based on color differences'. Proc. SPIE, Security and Watermarking of Multimedia Contents III, August 2001, vol. 4314, pp. 10–17.