

<http://ansinet.com/itj>

ITJ

ISSN 1812-5638

INFORMATION TECHNOLOGY JOURNAL

ANSI*net*

Asian Network for Scientific Information
308 Lasani Town, Sargodha Road, Faisalabad - Pakistan

Image Copy Detection Based on a Similar-sector Model and DFT

Ying Zheng and Yaping Lin

School of Information Science and Engineering, Hunan University,
No. 252, Lushan South Road, Changsha, 410082, China

Abstract: To protect authorized digital images from illegal use, image copy detection is very important for copyright protection. Its key issue is to extract robust feature to detect images with variety of attacks. In this study, a robust algorithm based on a similar-sector model and Discrete Fourier Transform (DFT) is proposed. An image is first divided by 48 concentric ellipses and 48 similar-sectors as the similar-sector model does, then 1D DFT magnitude spectrum which is obtained from the pixel gray values in each elliptical ring, is used as our image feature for copy detection. Our feature is not only robust to various signal modifications but also to geometric distortions. At last, the correlation coefficient is calculated to measure the similarity between the test image and the original image. The experimental results demonstrate that the proposed approach outperforms the existing methods and performs particularly well on detecting geometric distorted images, for instance, it can detect the image which is scaled, cut or rotated to any degree.

Key words: Image copy detection, a similar-sector model, feature extraction, robustness

INTRODUCTION

In the process of the fast spread of the network and multimedia technology, digital images can be easily duplicated and distributed. The infringement usage of them can result in seriously economic losses and legal dispute. Thus, it is a significant topic for protecting digital images from unauthorized use.

Generally, there are two approaches employed against copyright violation: watermarking (Jiang, 2012; Li *et al.*, 2011; Wu *et al.*, 2011) and Content-based Copy Detection (CBCD). Digital watermarking was the first way using for copyright protection. It needs to embed watermark which are visible or invisible into the images before they are distributed (Wolfgang and Delp, 1996). However the exiting techniques still lack robustness to a variety of attacks; and pirates can take some methods to remove a watermark. The key advantage of CBCD over watermarking is that it uses the image itself but no additional information to detect the copy. In detail, an image which is suspected as an unauthorized user using the content can be put into an image copy detector. The signature, extracted from the test image, is compared with the same signature which is extracted from each original image stored in the database of the copy detector to determine whether it is a copy or not. Consequently, CBCD can detect the illegal copies on its own and furthermore, it can complement watermarking as well (Wang *et al.*, 2010). One way to combine them is that the

copy detector returns the test image that is detected as a copy and then the owner can extract the watermark information in it to prove the ownership.

Strictly speaking, CBCD is a special category of Content-based Image Retrieval (CBIR) (Amsaleg and Gros, 2001; Zhu *et al.*, 2011; Huang *et al.*, 2012). However, for a test image, the latter is to retrieve images that are similar in color, shape or texture to it; the former aims to find its copies which may be suffered with signal or geometric attacks, such as saturation increased, blurring, Gaussian noising, rotation, scaling, etc. Therefore, if the existing CBIR technologies are applied to CBCD, it may cause the number of false alarms considerably.

Now-a-days, a number of approaches have been developed for image copy detection. Most exiting methods can fall into two classes: local feature-based (Kang *et al.*, 2011; Ke *et al.*, 2004) and global feature-based (Chang *et al.*, 1998; Kim, 2003; Wu *et al.*, 2006; Lin *et al.*, 2008; Li *et al.*, 2009). As we known, the SIFT (Kang *et al.*, 2011) and PCA-SIFT (Ke *et al.*, 2004) descriptors were widely designed as local features in the field of image copy detection. Nevertheless, since the original SIFT regions are too small and short of completeness, the SIFT descriptors will be ambiguities. On the other hand, they are impressionable to signal modifications, such as noise. So the local features are not so suitable for image copy detection. The global features have a good description of integrity and structure. (Chang *et al.*, 1998) was the first to propose a near-replica

search engine called RIME (Replicated IMage dETector) for detecting unauthorized copies of images on the Internet. The authors used wavelets and C1C2C3 color space to characterize images. Although the method can detect slightly modified images with high accuracy, it is invalid to detect seriously distorted images. To increase the features' robustness and discrimination (Kim, 2003) used ordinal measure of DCT coefficients and employed the MAP (maximum a posteriori) criterion to find an optimal threshold. Kim's method can only discover copies with rotation by 180° or very small angles, fails to resist other angles. It can neither detect the cropped images. Wu *et al.* (2006) proposed an elliptical track division strategy to extract two kinds of features to overcome the defect of Kim's approach. However, when the degree of rotation is larger than 10° (except 180 and 270°), the method cannot detect. Moreover, the method is not effective to find the copies having some general signal manipulations, such as water coloring and mosaic tile. Lin *et al.* (2008), firstly extracted the edge feature from an image and then a center detection mechanism and two clustering approaches were combined. The method performs well on detecting general modified images as the edge is not sensitive to the majority of distortions. Since the regions which they selected to compute the pixel features are rectangular, it can only detect the rotated images with the degree of one fourth of 90° , two fourths of 90° , three fourths of 90° and so on. Li *et al.* (2009) proposed a scheme based on the Gabor texture descriptor. They used Gabor filters and 2D-DFT. Then, a circular shift method is employed on the scale invariant feature to achieve the rotation invariance. Their method is robust to small angles rotation (e.g., $1, 5, 13^\circ$ etc.) and some special rotation (e.g., $90, 180, 270^\circ$ etc.). But for larger angles of rotation (e.g., $40, 60, 80^\circ$ etc.), it has a high false alarm rate. Observe that, according to the above interpretation, the recent global image features are effective and discriminative for image copy detection but their key issue is to be vulnerable to rotation transformation. As one dimension Discrete Fourier Transform (1D DFT) has a property of shifting invariance, the problem of rotation is transformed into shifting. In the study, a robust method based on a similar-sector model and DFT is proposed for image copy detection. The main contributions of this study are: (1) A similar-sector model is built, (2) The DFT-based descriptor is proposed. Even if the copy is scaled, cut or rotated to any degree, the extracted features are nearly the same as those of its original image and (3) A correlation coefficient is calculated to measure the similarity between the query image and the test image.

THE PROPOSED SCHEME

Here, a similar-sector model and the feature extraction method are introduced and

then the robustness of the proposed method to rotation is theoretically analysed.

A similar-sector model: The similar-sector model describes a strategy of how to divide an image. Because the shape of an image is rectangle, the image is regarded as a rectangle template with length a and width b . Firstly m ellipses are drawn in the rectangle template. One ellipse is the inscribed ellipse of the rectangle, that is to say, its Semiminor axis and Minor axis are $a/2, b/2$, respectively. All the ellipses are concentric and the difference between the Semiminor Axes of two adjacent ellipses is a constant a/m (Fig. 1).

In Fig. 2, the ellipse is divided across its center into n similar-sector regions, each with an equal angle of $2\pi/n$. As once the pixel location is out of the biggest ellipse, it is ignored. In practice, they are the corner of the image and are not so important to an image. So it has a negligible influence on image copy detection.

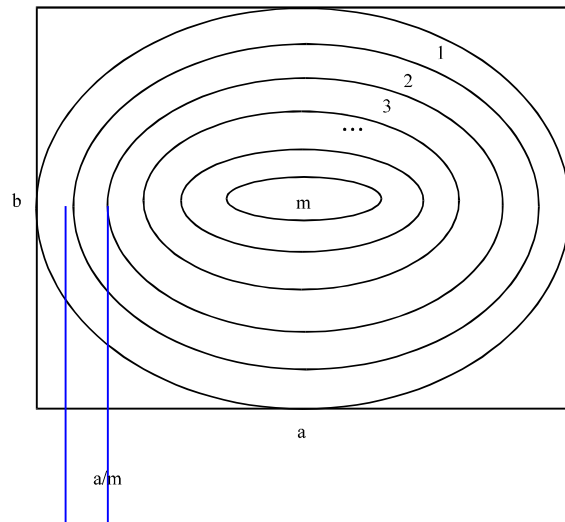


Fig. 1: m ellipses, m : No. of ellipses

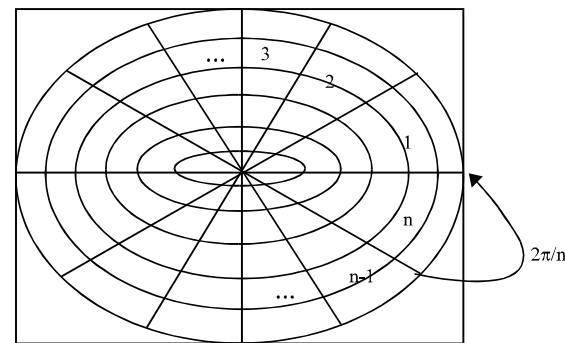


Fig. 2: A similar-sector model, m : No. of similar-sector blocks

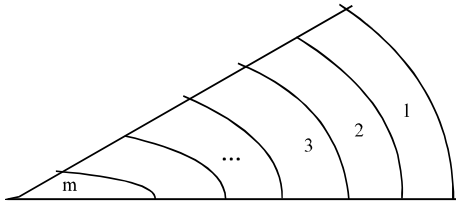


Fig. 3: m similar-sector blocks, m: No. of similar-sector blocks

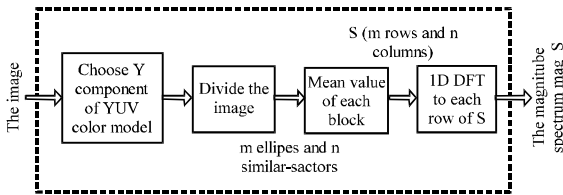


Fig. 4: Flowchart of the procedure of the image's feature extraction

Three concepts are defined as follows:

- **The elliptical ring:** The area between two adjacent ellipses is called an elliptical ring
- **The similar-sector block:** The area in a similar-sector region between two adjacent ellipses is called a similar-sector block (Fig. 3)
- **The gray value:** The gray value in each similar-sector block is presented by the average gray pixel value of the block

As the image is rotated with a certain angle, the similar-sector rings in each ellipse ring will rotate in the same way. Although, there are some fluctuations between the pixels in the corresponding similar-sector ring of two images, the difference was not significant so the mean values of two similar-sector rings are nearly identical.

The procedure of feature extraction: 1D DFT has a characteristic of shifting invariance. A time shift x_0 of signal is equivalent to a phase shift ωx_0 in the frequency domain and the magnitude remains unaltered (Oppenheim *et al.*, 1999). So, if the rotate of an image can be transformed into the shifting of arrays, the DFT-based features will be robust to rotation. The detail proof of the proposed feature that is robust to rotation is described earlier. Here, the procedure of feature extraction is introduced.

Figure 4 is the flowchart of the feature extraction procedure. If the image is color, it is first converted into YUV color model. Then only the Y component is used in the scheme.



Fig. 5: Dividing the image

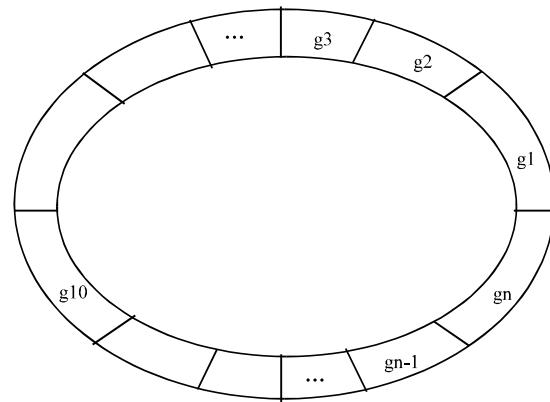


Fig. 6: The i th ellipse ring, g_x : Gray value of the x th ($x = 1, 2, \dots, n$) similar-sector block

The image is divided with m ellipses and n similar-sectors the way as a similar-sector model does (Fig. 5). One similar-sector region contains m similar-sector blocks. Each elliptical ring has n gray values. Figure 6 shows the i th ($i = 1, 2, \dots, m$) ellipse ring with n blocks, each value of the blocks is stored into an n-length array x , where $x(1) = g_1, x(2) = g_2, \dots, x(n) = g_n$. Then the 1D DFT magnitude spectrum mag_x of the array x is got. For a given image, as m ellipses is used to divide, m arrays are obtained. The arrays are stored into S (m rows, n columns). Then the 1D DFT magnitude spectrums of each row of S are stored into mag_S (m rows, n columns). Finally, the magnitude spectrums mag_S are regarded as the image's signature. The algorithm is shown in Table 1 in detail.

In the study, the correlation coefficient between two images' feature vectors is used to measure the similarity. If given an original image, the magnitude spectrum mag_S with m rows and n columns is obtained, let mag_S^i denote the i th ($i = 1, 2, \dots, m$) row of mag_S ; given a test image, the magnitude spectrum mag_T with m rows n columns is

Table 1: Feature extraction algorithm

```

Algorithm: Feature extraction
//The number of ellipses and similar-sectors: m and n
// Array picmg stores the pixel value in an ellipse block
// Array s[m, n] stores the gray values in ellipse blocks, s(k,:) is the kth row of s
//fft(u) is a function that can calculate the 1D-DFT of u
//Array f[m, n] stores the 1D-DFT to s, f(k,:) is the kth row of f
//abs(u) is a function that can calculate the magnitude spectrum of u
//mag_S is the magnitude spectrum
For x is each ellipse
For y is each similar-sector
picmg stores each pixel (i, j) in the similar-sector ring
s(k,:) is the average value of the picmg
f(k,:) = fft(s(k,:))
mag_S = abs(s(k,:))
end
end
Output mag_S
    
```

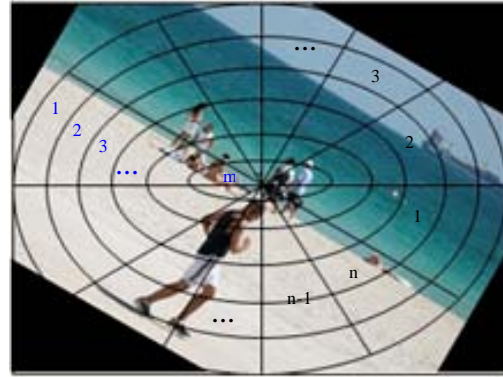


Fig. 7: Dividing the rotated image

obtained, let mag_T^i denotes the i th row of mag_T . Then the correlation coefficient of mag_S^i and mag_T^i is calculated. The formula is described as follows:

$$r_i = \frac{\sum_{j=1}^n (mag_S_j^i - \overline{mag_S^i})(mag_T_j^i - \overline{mag_T^i})}{\sqrt{\sum_{j=1}^n (mag_S_j^i - \overline{mag_S^i})^2} \sqrt{\sum_{j=1}^n (mag_T_j^i - \overline{mag_T^i})^2}} \quad (1)$$

Where:

$$\overline{mag_S^i} = \frac{1}{n} \sum_{j=1}^n mag_S_j^i \quad (2)$$

$$\overline{mag_T^i} = \frac{1}{n} \sum_{j=1}^n mag_T_j^i \quad (3)$$

So the average m of correlation coefficients is:

$$\bar{r} = \frac{1}{m} \sum_{i=1}^m r_i \quad (4)$$

Given a preset threshold T , if $\bar{r} > T$, then the test image is the copied vision of the original image, else it is not a copy.

Robust against rotation: It can be proved that the extracted feature is robust to rotation. Figure 7 shows a test image which is the original image in Fig. 6 rotated clockwise with 30° . It is also divided by m ellipses and n similar-sectors. As described in the similar-sector model, the rotation of the image means the rotation of each ellipse ring. Consequently, it will result in a shift of the array x . Figure 8 shows the i th ($i = 1, 2, \dots, m$) ellipse ring of the rotated image with n blocks, each value of the blocks is stored into an n -length array y , where, $y(1) = g_2, y(2) = g_3, \dots, y(n-1) = g_n, y(n) = g_1$.

As discussed in subsection 2.2.2, the 1D DFT magnitude spectrum mag_x of the array x ($x(1) = g_1$,

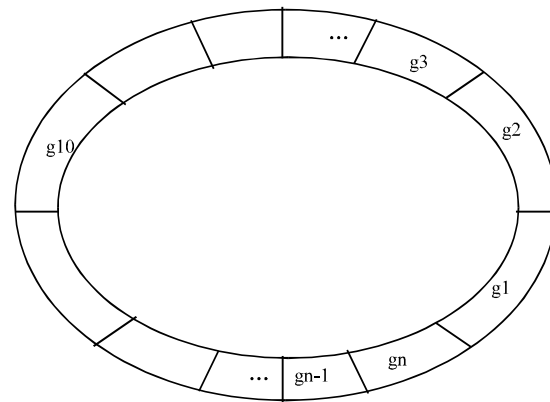


Fig. 8: The i th ellipse ring

$x(2) = g_2, \dots, x(n) = g_n$) is obtained. Now the 1D DFT magnitude spectrum mag_y of the array y is got. According to the shift invariance of the Fourier transform, mag_x is equivalent to mag_y . Likewise, the magnitude spectrum obtained from the i th ($j = 1, 2, \dots, m, j \neq i$) ellipse ring of the original image is the same as it from the test image. Therefore, the DFT magnitude spectrum as the image's feature is robust to rotation.

EXPERIMENTS AND ANALYSIS

Image dataset: To evaluate the performance of the proposed scheme, the UCID (Uncompressed Color Image Database) dataset (Schaefer and Stich, 2004) is used which contains 1,338 color images in sizes of 512×384 or 384×512 . In the experiment, 20 images are randomly selected from the dataset (the original dataset), showed in Fig. 9. The selected images have gone through 29 attacks using Photoshop software. Figure 9 shows the 29 processed images of Fig. 9 (1-17) are the images which are geometric attacked and Fig. 9(18-30) are the images which suffer signal attacked. In total, 580 images are obtained in the copied database (the copied dataset).

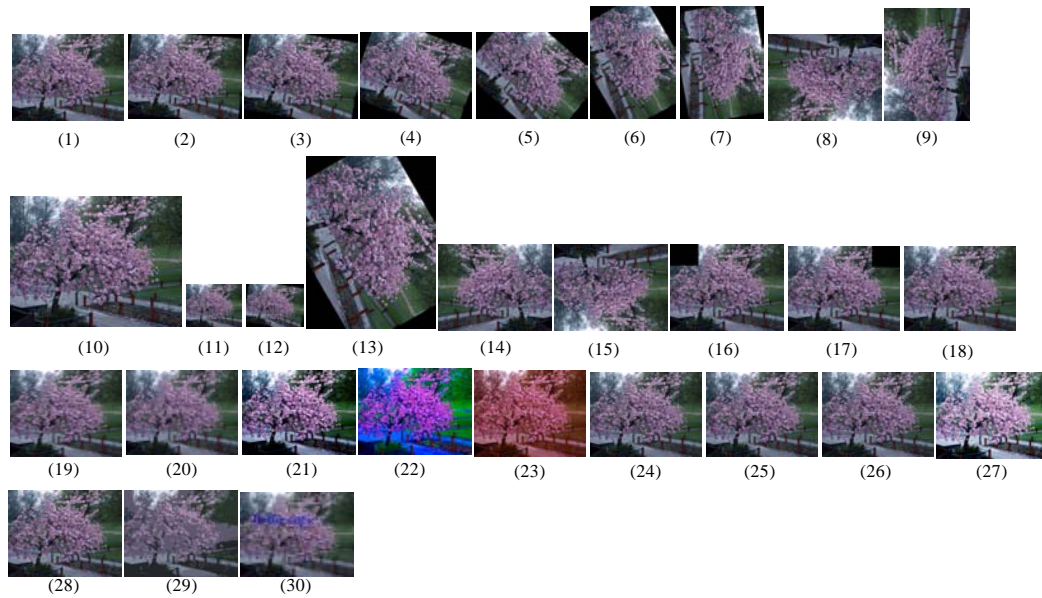


Fig. 9: An original image and its 29 variations, (1) Original image, (2) Rotation (5°), (3) Rotation (10°), (4) Rotation (20°), (5) Rotation (40°), (6) Rotation (60°), (7) Rotation (80°), (8) Rotation (180°), (9) Rotation (270°), (10) Scaling (150%), (11) Scaling (50%), (12) Rotation (10) Scaling (50%), (13) Rotation (60) Scaling (150%), (14) Horizontal flipping, (15) Vertical flipping, (16) Cut left upper, (17) Cut right upper, (18) Blurring, (19) Motion blurring, (20) Mosaic, (21) contrast enhancement, (22) Saturation increased, (23) Hue change, (24) Sharpening, (25) Splash, (26) Veining, (27) Histogram equalization, (28) Gaussian noising (12.5%), (29) Woodcut and (30) Inserting text

Evaluation criteria: To evaluate the performance of the proposed scheme, the precision and recall curve (PR curve) are used. Let N_T be the total number of the copied images; N_c the number of the copied images whose respective correlation coefficients are larger than T ; N_{Nc} the number of the non-copied images whose respective correlation coefficients are larger than T . The PR curve is defined as:

$$\text{Precision} = N_c / (N_c + N_{Nc}) \quad (5)$$

$$\text{Recall} = N_c / N_T \quad (6)$$

So PR curves can be generated by selecting different values of threshold T (the X-axis denotes the recall and the Y-axis denotes the precision). A higher value of precision rate means a higher accuracy and a lower rate of false alarms and a higher value of recall rate means minimizing the number of missed copies, as is evident from the definitions. Obviously, the higher the precision and recall rate is, the better the performance is.

Experiment and analysis: In this subsection, the first experiment is designed to select the appropriate

number of the ellipses and similar-sectors. Then two experiments are made to evaluate three methods: (Kim, 2003; Wu *et al.*, 2006) and the proposed method. Kim used ordinal measure of DCT coefficients, Wu *et al.* (2006) proposed an elliptical track division strategy and computes the mean and variance value of each elliptical track.

The number of ellipses and similar-sectors: As the performance of the proposed scheme depends on the number of the ellipses and similar-sectors which are used, the first experiment is conducted to select the appropriate number by using the original dataset and the copied dataset. Assume that the number of the ellipses and similar-sectors is (x, y) . (x, y) is successively set to (8, 8), (16, 16), (32, 32), (48, 48), (64, 64). As Fig. 10 shows, when both x and y are 48 or larger, the curve is very close to the coordinate point (1, 1). Namely, almost all the copied images can be detected successfully. And when the number is larger, the accuracy rate remains nearly the same but the running time is much longer. Therefore consider both of the accuracy and the running time, 48 ellipses and 48 similar-sectors are used in the remaining experiments.

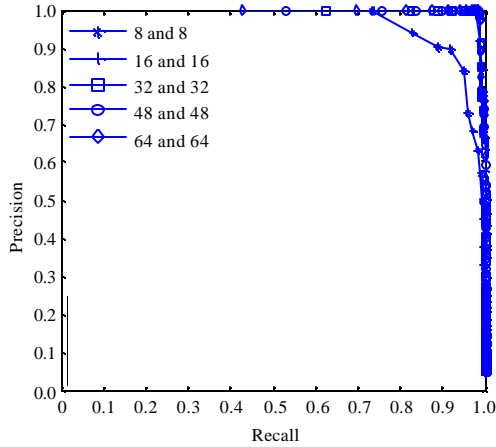


Fig. 10: PR curves for the different number of the ellipses and similar-sectors

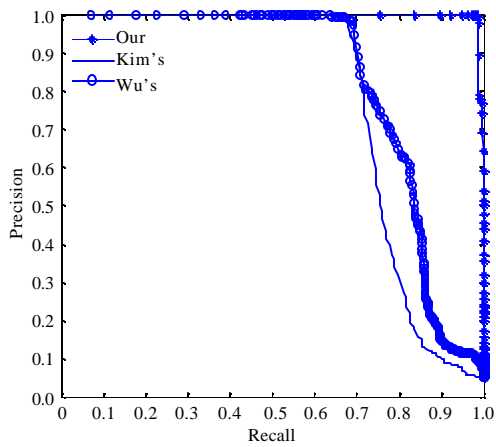


Fig. 11: PR curve for all attacks

Comparing three methods: The second experiment is made to evaluate the performance for all 29 attacks. The original dataset and the copied dataset are used. The average precision and recall curves for the three methods of Kim's, Wu's and the proposed approach are showed in Fig. 11. At the recall of 90%, the proposed approach is the best with a precision of 100%, Kim's and Wu's are 14 and 10%, respectively. At the recall 70%, the precision of Kim's and Wu's are both 90%, the proposed scheme achieves 100% precision again. Hence, the proposed method is the best among the three on the detection performance.

The third experiment shows the detecting ability of Kim's, Wu's and the proposed approach more in detail. It is based on the UCID dataset and the copied dataset. The threshold is set according to the P-R curve in the second

Table 2: Successful detection results for 580 modified images

Attack	Kim's (Kim, 2003)	Wu's (Wu <i>et al.</i> , 2006)	Proposed approach
Rotation (5°)	15	19	20
Rotation (10°)	5	2	20
Rotation (20°)	0	1	20
Rotation (40°)	0	0	20
Rotation (60°)	0	0	20
Rotation (80°)	0	1	20
Rotation (180°)	20	20	20
Rotation (270°)	0	20	20
Scaling (50%)	20	20	20
Scaling (150%)	20	20	20
Rotation+scaling(10°+50%)	5	2	20
Rotation+scaling(60°+200%)	0	0	20
Horizontal flipping	20	20	20
Vertical flipping	20	20	20
Cut left 25%	9	9	20
Cut right 25%	12	11	20
Blurring	20	20	20
Motion blurring	20	20	20
Mosaic	20	20	20
Contrast enhancement	20	12	17
Saturation increased	18	13	20
Hue change	20	20	20
Sharpening	20	20	20
Splash	20	20	20
Veining	20	20	20
Histogram equalization	20	13	16
Gaussian noising (12.5%)	20	20	20
Woodcut	18	20	20
Inserting text	19	20	20

experiment. Actually, the threshold can be adjusted to satisfy the special requirements. For example, some institutions require a high recall ratio, the precision ratio is not yet so important. In the experiment, the precision and recall are balanced and they are optimal. Hence, the threshold of the proposed approach is 0.9880, Kim's is 0.4530 and Wu's is 0.0750. Namely, in the proposed approach, if the correlation coefficient of the two images is larger than the threshold, the test image is a copied vision of the original image. While in Kim's and Wu's methods, when two images' distance is smaller than the threshold, the test image is treated as a copy. Table 2 column 1 shows a variety of modifications of the images; Column 2~4 show the number of successful detections for each scheme under each attack, respectively.

As 20 images are chosen that are attacked, the value of a cell in the Table 2 must be 20 if every original image can be detected. From Table 2, Kim's method can detect some rotated images with small angles or 180° but it performs ineffective to detect the image which is rotated with large angles, such as 20, 40, 60, 80 and 270°. When the image is cut away some pixels like Fig. 9(16-17), Kim's method is not desirable either. Wu's method is robust against the rotation by 180, 270° or small angles but vulnerable to large angles. It also cannot detect all the contrast enhancive images, saturation increased images or histogram equalized images. As to the proposed scheme, it can detect the images rotated to any angles. It

is robust to rotation, scaling, flipping, text-inserting, noising, sharpening and so on. The results confirm that the detection performance on signal modifications of the proposed approach behaves mainly the same as Kim's and a little better than Wu's. However, the proposed approach shows an obvious advantage on geometric modifications such as rotation and cutting.

CONCLUSION

In this study, a 1D-DFT descriptor is used to detect the copied versions of images. It is evident that the proposed method is able to detect the images which are undergone with any angles of rotation attacks but all the other methods based on global features can only detect those rotated with specific angles. Moreover, our method can also detect the cropping and scaling images and it is robust to general signal modifications. The implementation algorithm of the proposed method is relatively simple, so it is also suitable for the copy detection of large scale images. As the method extracts the global feature, it has limitations on detecting an image which is a small part of the original image. So in the future, extending the proposed method to a detector that combines with extracting the global and local features may be a topic that is worth further study.

ACKNOWLEDGMENTS

This study was partially supported by the Ph.D. programs foundation of Ministry of Education of China (20100161110025).

REFERENCES

- Amsaleg, L. and P. Gros, 2001. Content-based retrieval using local descriptors: Problems and issues from a database perspective. *Pattern Anal. Appl.*, 4: 108-124.
- Chang, E.Y., J.Z. Wang, C. Li and G. Wiederhold, 1998. RIME: A replicated image detector for the world-wide web. *Proceedings of the SPIE Multimedia Storage and Archiving Systems III*, November 2-5, 1998, Boston, USA., pp: 58-67.
- Huang, Z.L., S.Z. Guo and Z.M. Lu, 2012. Image retrieval based on topological features of gray-level co-occurrence networks. *Inform. Technol. J.*, 11: 626-631.
- Jiang, M.F., 2012. A new adaptive visible watermarking algorithm for document images. *Inform. Technol. J.*, 11: 1322-1326.
- Kang, L.W., C.Y. Hsu, H.W. Chen, C.S. Lu, C.Y. Lin and S.C. Pei, 2011. Feature-based sparse representation for image similarity assessment. *IEEE Trans. Multimedia*, 13: 1019-1030.
- Ke, Y., R. Sukthankar and L. Huston, 2004. Efficient near-duplicate detection and sub-image retrieval. *Proceedings of the ACM International Conference on Multimedia*, October 10-16, 2004, New York, USA., pp: 869-876.
- Kim, C., 2003. Content-based image copy detection. *Signal Process.: Image Commun.*, 18: 169-184.
- Li, J., R.D. Wang and J. Zhu, 2011. A watermark for authenticating the integrity of audio aggregation based on vector sharing scheme. *Inform. Technol. J.*, 10: 1001-1008.
- Li, Z., G. Liu, H. Jiang and X. Qian, 2009. Image copy detection using a robust Gabor texture descriptor. *Proceedings of the 1st ACM Workshop on Large-Scale Multimedia Retrieval and Mining*, October 19-24, 2009, Beijing, China, pp: 65-72.
- Lin, C.C., N. Klara and C.J. Hung, 2008. An image copy detection scheme based on edge features. *Proceedings of the IEEE International Conference on Multimedia and Expo*, June 23-26, 2008, Hannover, Germany, pp: 665-668.
- Oppenheim, A.V., A.S. Willsky and S.H. Nawab, 1999. *Signals and Systems*. 2nd Edn., Prentice Hall International Inc., China, ISBN-13: 9787302030584, Pages: 983.
- Schaefer, G. and M. Stich, 2004. UCID: An uncompressed color image database. *Proceedings of the SPIE Storage and Retrieval Methods and Applications for Multimedia*, January 20, 2004, San Jose, CA., USA., pp: 472-480.
- Wang, Y.G., Y. Lei and J. Huang, 2010. An image copy detection scheme based on radon transform. *Proceedings of the 17th IEEE International Conference on Image Processing*, September 26-29, 2010, Hong Kong, China, pp: 1009-1012.
- Wolfgang, R.B. and E.J. Delp, 1996. A watermark for digital images. *Proceedings of the IEEE International Conference on Images Processing*, Volume 3, September 16-19, 1996, Lausanne, Switzerland, pp: 219-222.
- Wu, C.H., Y. Zheng, W.H. Ip, Z.M. Lu, C.Y. Chan and K.L. Yung, 2011. Effective hill climbing algorithm for optimality of robust watermarking in digital images. *Inform. Technol. J.*, 10: 246-256.
- Wu, M.N., C.C. Lin and C.C. Chang, 2006. A robust content-based copy detection scheme. *Fundam. Inform.*, 71: 351-366.
- Zhu, J., S.C.H. Hoi, M.R. Lyu and S. Yan, 2011. Near-duplicate keyframe retrieval by semi-supervised learning and nonrigid image matching. *J. ACM Trans. Multimedia Comput. Commun. Appl.*, Vol. 7.