Journal of Information Hiding and Multimedia Signal Processing ©2016 ISSN 2073-4212
Ubiquitous International Volume **7**, Number **5**, September **2016**

A Robust Image Copy Detection Method Based On Feature Extraction Algorithm

Jianfeng Lu[1], Xiaoyan Xu [2], Li Li[1*], Chin-Chen Chang[3*], Qiang Wang[2]

[1]Institute of Graphics and Image, Hangzhou Dianzi University, Hangzhou, Zhejiang China
[2]Institute of Digital Media, Hangzhou Dianzi University, Hangzhou, Zhejiang China
[3]Department of Information Engineering and Computer Science, Feng Chia University, Taichung, Taiwan, alan3c@gmail.com
*Corresponding author: Chin-Chen Chang and Li Li

Received January, 2016; revised May, 2016

Abstract. *Detecting copied versions of a query image is very important for copyright protection. This paper presents a robust copy detection method for images. The main contribution of our algorithm is to combine the image feature extraction and the saliency image segmentation methods, and a novel two steps approach is proposed for the image copy detection. Firstly, several salient regions in a query image are extracted based on salient region detection method and segmented into a series of partial images, with introduction of illumination variation to obtain a better saliency image. Secondly, surf key points in both partial images and images in database are extracted and the number of correctly matched key points is computed. Similarity value is computed to judge if two key points are correctly matched. Finally, if the number of correctly-matched key points is greater than the threshold value, the image from database is judged as a copy of the query image. The experimental results show that the proposed method can detect copyright information of query image even when the image is cropped, rotated, put in picture, water-colored, changed in contrast, blurred or inserted with words.*

Keywords: Image copy detection; Salient region detection; Key points matching.

1. **Introduction.** Digital image transmission is becoming much easier and simpler with development of multimedia technology. However, simplicity of digital image transmission obstructs the protection of copyrights of digital image, which causes copyright owners enormous loss. Therefore, it is important to test whether an authorship of image is copied and given away illegally.

Image copy detection is a technology that can appraise copyright of image. A database is established in copy detection system. If the owner of an image suspects his/her image is being used illegally, he/she can raise a query to the copy detection system. There are two main methods to detect copies at present, one is based on global features of image and the other is based on local features. In copy detection methods based on global features of the original image [1, 2, 3, 4], a series of 8*8 blocks were segmented and feature of every block was extracted by Discrete Cosine Transformation(DCT). However, these methods are in inferior robustness when resisting cropping and picture in picture transforms. For copy detection methods based on local features[5, 6, 7], most of them consist of three steps: (1) keypoint detection; (2) keypoint patch normalization; and (3) descriptor computation. After calculating descriptor, keypoints between two images needed to be matched, and mis-matched keypoints can be removed through RANSAC(Random Sample Consensus)[8, 12].

Transforms for copy detection is shown in Table 1. Cropping, picture in picture and text insertion are difficult to solve. Picture in picture refers to cutting a part of original image and implanting it into another image. At the moment, there are few papers discussing how to resist transforms above. Lin et al.[6] put forward a copy detection method based on features at image edge which could resist most transforms above. While this paper merely discussed solution of transform on image stitching but not on truly picture in picture. Analyzing cropping and picture in picture, it can be found that these two transforms cut unimportant parts of the original image while keep important parts. Thus important parts of original image should be used to solve the two transforms.

In this paper, a robust copy detection algorithm is proposed. Firstly, instead of matching query image with images in database, a series of partial images are used to match. These partial images are segmented from query image based on salient region detection. Secondly, surf key points in every partial image and in images from database are extracted, respectively. The number of correctly-matched key points is computed based on the proposed key point matching method. Finally, if the number of correctly-matched key points is greater than the threshold value, the image in database is judged as a copy of query image. The experimental results show that the proposed method can resist most transforms listed in Table 1 with a superior detection ratio.

Table 1. List of transforms for copy detection

Transform No.	Description	Transform No.	Description
C1	Picture in picture	C9	Text insertion
C2	JPEG compression	C10	Rotation
C3	Change of color	C11	Resizing
C4	Change of illumination	C12	Decrease in image quality
C5	Change of contract	C13	Gaussian noising
C6	Cropping	C14	Water coloring
C7	Blurring	C15	Mosaic tile
C8	Image flipping		

The paper is organized as follows. In Section 2, the major steps of the proposed method are introduced. Then the new salient region detection and key points matching method are provided. Experimental results are analyzed in Section 3 and the conclusion is drawn in Section 4.

2. Proposed robust image copy detection method.

2.1. **Introduction of the proposed method.** The main advantage of this method is robustness. The method is based on local keypoints and salient region detection. The technique of salient region detection is used to segment the query image into partial images. Each partial image contains one salient region of query image. The technique of local keypoints is used to match partial images with images in the database. Surf keypoints are chosen for its good performance under illumination change, blurring, and geometric transforms. The flow diagram of the method is shown in Figure 1, where, $N\_matched$ is the number of the correctly matched key points, T is the threshold value.

The method is demonstrated in Figure 1 and summarized as follows:

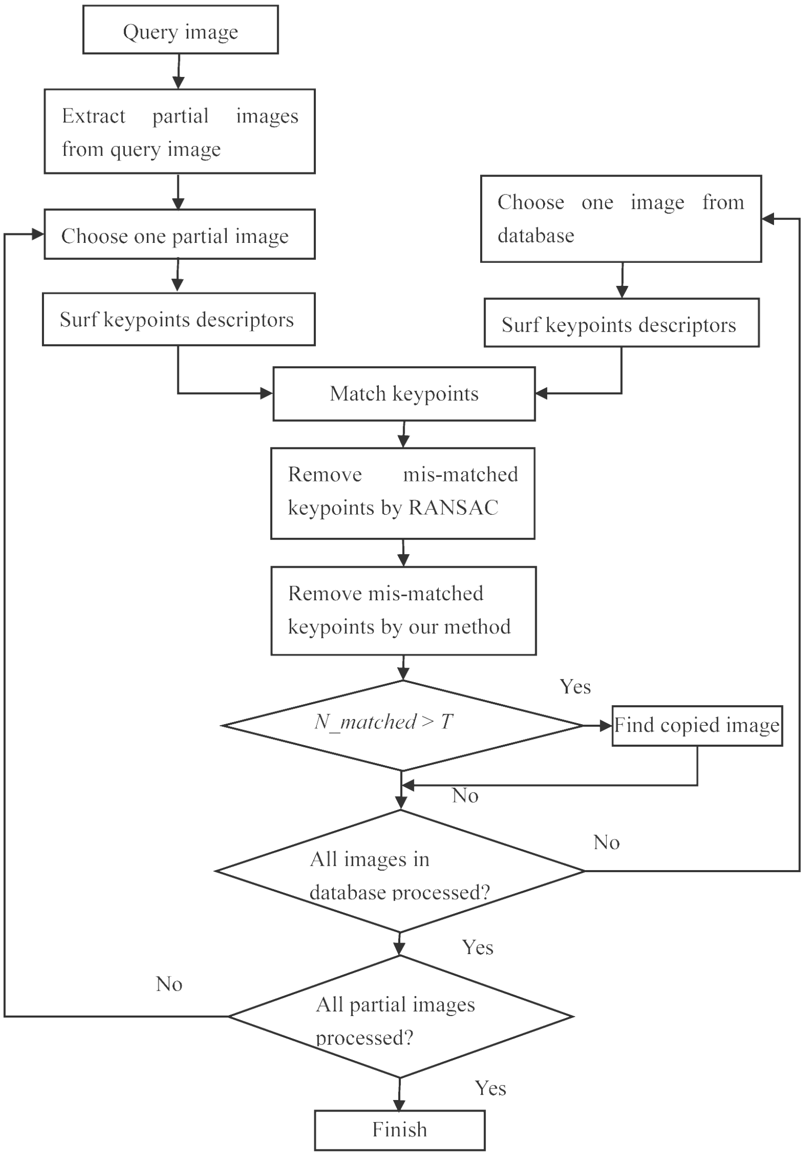


FIGURE 1. The flow diagram of the proposed method

Step1: Segment query image into a series of partial images based on the proposed salient region detection.

Step2: Choose one partial image and choose one image from database.

Step3: Match partial image with the image from database by surf keypoints.

Step4: Remove mis-matched keypoints by RANSAC.

Step5: Remove mis-matched keypoints by the proposed keypoints matching.

Step6: If the number of correctly matched keypoints are greater than the threshold value, the image in database is judged as a copy of query image.

Step7: If any partial image has not been processed with the images from database, go to Step 2.

In the proposed method, segmentation of query image into a series of partial images and keypoints matching are the two important procedures. As shown in Cheng et al.'s method[10] and Perazzi et al.'s method[11], there are four steps in salient region detection:

(1) Decompose the image into basic elements that preserve relevant structure. (2) Compute each element's saliency value. (3) Assign saliency values to every pixel. (4) Obtain the high quality profile of saliency image. Out of the four steps above, the second and the fourth steps are improved in this paper. In the second step, illumination variation is added into saliency value computation to obtain a better salient image. In the fourth step, a self-adaptive threshold is computed. The proposed salient region detection method is introduced in Subsection 2.2.

As for matching keypoints, RANSAC is used to remove mis-matched keypoints. However, it doesn't always work well. In this paper, two areas around the matched keypoints are extracted correspondingly and then their structural similarity value is computed in RGB color space to judge if these two keypoints are correctly matched. The proposed keypoints matching method is introduced in Subsection 2.3.

2.2. **Proposed salient region detection.** The saliecy value computation in this paper is improved based on Cheng et al.'s method[10]. Saliecy value computation in Cheng et al.'s method is defined as:

$$S(E_k) = w_s(E_k) \sum_{r_k \neq r_i} e^{\frac{D_s(E_k,E_i)}{-\sigma_s^2}} w(E_i) D_E(E_k, E_i) \tag{1}$$

where, $D_s(E_k,E_i)$ is the spatial distance between elements E_i and E_k. $D_E(E_k, E_i)$ is the color distance metric between elements $E_k$and E_i . σ_s^2 controls the strength of spatial distance weighting, and σ_s^2=0.4. w(E_i) is the weight of region E_i defined by the number of pixels in E_i, and $w_s(E_k)$ is a spatial prior weighting term similar to center bias.

As shown in Eq.(1), three factors have been considered in salient region detection: color difference between elements, a weight defined by the number of pixels in the element and a weight defined by elements spatial distance. However, every element has its own feature. The connection area among objects is known to be the background area with similar color. The background is regarded as non-salient region. On the other hand, the area contains an object is regarded as salient region. Based on the above discussion, illumination variation of every element is added. If an element is in non-salient region, its illumination variation will be smaller. But if an element is belong to a salient region, its illumination variation will be larger. The k-th elements illumination variation is defined as D_l^k , as shown in Eq.(2).

$$D_E^k = \sum_{i=1}^{N(E_k)} d_E^i / N(E_k) \tag{2}$$

where, N(E$_k$) is the number of pixels contained of *k-th* element, d_i^E is the convolution of 3*3 neighborhood of pixel i and template $\begin{pmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{pmatrix}$, and k is the order of element.

Furthermore, the locations of salient regions are not always in the center of image, so in the proposed equation, the spatial prior weight is not included. The proposed salient region detection equation is defined as:

$$S(E_k) = D_E^k \sum_{E_k \neq E_i} e^{\frac{D_s(E_k,E_i)}{-\sigma_s^2}} w(E_i) D_E(E_k, E_i) \tag{3}$$

The improved saliency images are shown in Figure 2. Compared with the other methods, the saliency images obtained by the proposed method contains darker non-salient regions and brighter salient regions. It is helpful to segment query image into a series of partial images.

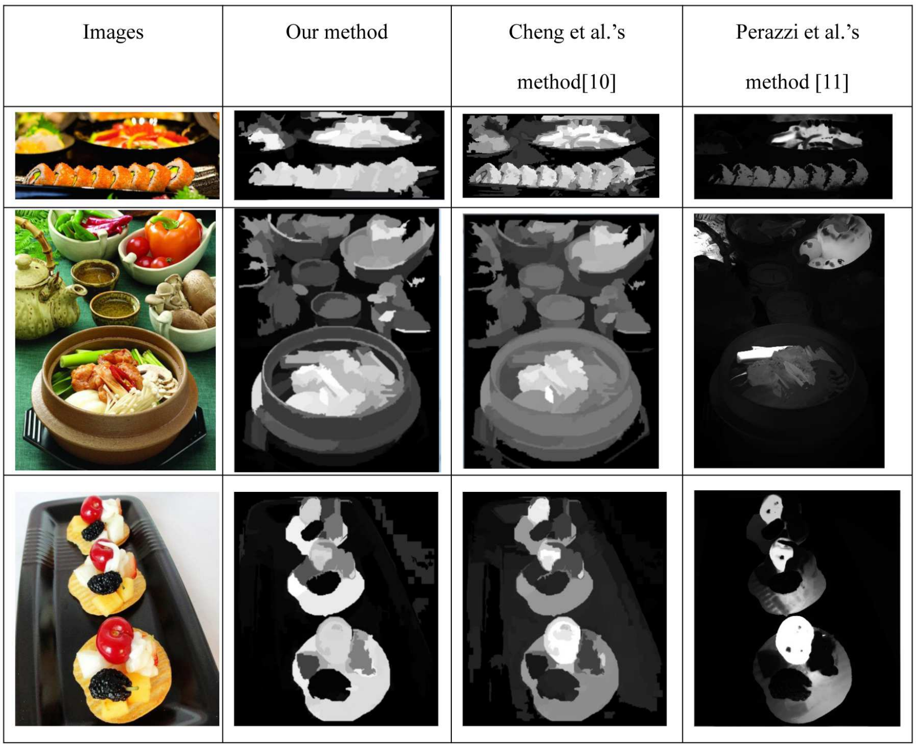


FIGURE 2. Saliency images of different methods

In the part of obtaining high quality profile, a fixed threshold 70 was used in Cheng et al.'s method. In this paper, average saliency value is computed, which is self-adaptive. It is defined as:

$$T = \frac{\sum_{i=1}^{n} S(i)}{n} \tag{4}$$

where, n is the number of pixels in the image, $S(i)$ is the salient value of each pixel in the image.

After obtaining a high quality profile, the minimum enclosing rectangle of each object is computed and the query image is segmented into a series of partial images. Each partial image contains an object. The partial images are used to match with images in the database. Figure 3 shows some examples of partial images.

2.3. **Proposed key points matching algorithm.** After removing the mis-matched keypoints by RANSAC, it should be further determined whether the remaining keypoints are correctly matched. A pair of remaining keypoints (Q_i, I_j) is chosen from one partial image of query image Q and image I. A 5*5 square region centered around the keypoint, and oriented along the keypoint's orientation is extracted. The square block is in the coordinate system and the scale of the keypoint. Thus, two blocks $Block\_Q_i$ and $Block\_I_j$ are extracted. Their structural similarity value in RGB color space is computed as shown in Eq.(5).

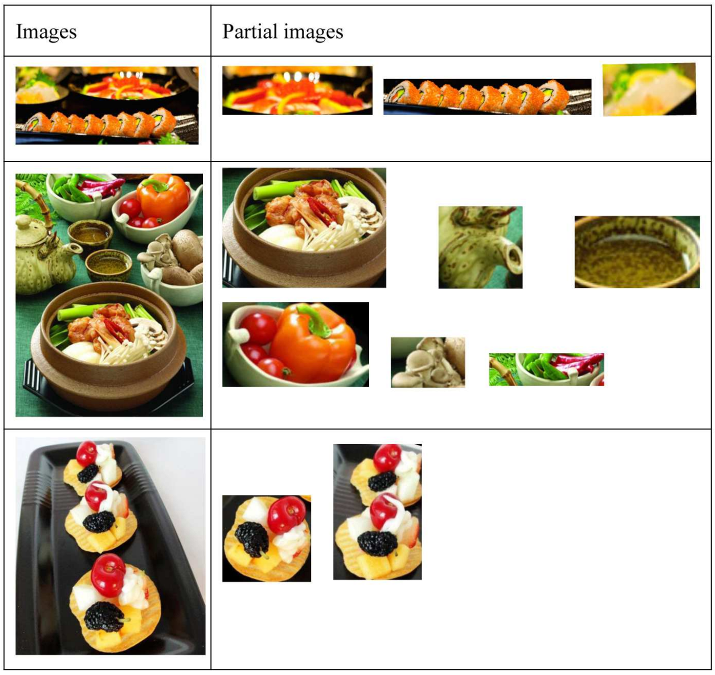


FIGURE 3. Query images and their partial images

$$(ssim\_r_i, ssim\_g_i, ssim\_b_i) = SSIM(Block\_Q_i, Block\_I_j) \tag{5}$$

The symbol $ssim\_r_i$ is defined as the structural similarity value in red channel,$ssim\_g_i$ is the structural similarity value in green channel and $ssim\_b_i$ is the structural similarity value in blue channel.SSIM[11] is a method to measure the similarity of two images. The high value means the two images are similar.

The final structural similarity value $ssim\_c$ between $Block\_Q_i$ and $Block\_I_j$ is obtained from Eq.(6). Based on a large number of experiments, the matching is significantly correct when $ssim\_c$ is greater than 0.55.

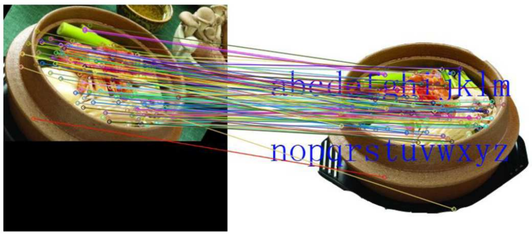
$$ssim\_c = \begin{cases} ssim\_r, & Q_i\_r > Q_i\_g and Q_i\_r > Q_i\_b \\ ssim\_g, & Q_i\_g > Q_i\_r and Q_i\_g > Q_i\_b \\ ssim\_b, & Q_i\_b > Q_i\_r and Q_i\_b > Q_i\_g \end{cases} \tag{6}$$

where, Q_i_r is the red value of Q_i, Q_i_g is the green value of Q_i, and Q_i_g is the blue value of Q_i.

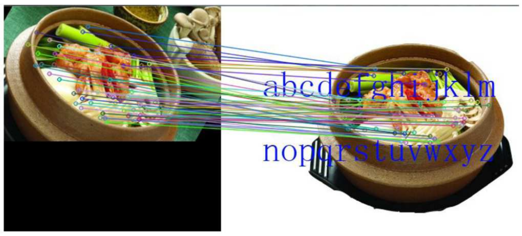
In Figure 4(a), there are some mis-matched keypoints left after RANSAC. As shown in Figure 4(b), the proposed keypoint matching method can keep the correctly-matched keypoint and remove the mis-matched key points.

To illustrate the performance of the proposed method further, the distinction of the proposed method and RANSAC is computed. A good matching method should keep more correctly matched key points between original image and its transformed images than between original image and its similar image.

In Figure 5(a), the correctly matched key points number is greater than 10 when matching original image and its transformed image. In the process of matching original image and its similar image, the correctly matched key points number is much smaller than 10.



(a) Remove mis-matched keypoints by RANSAC



(b) Remove mis-matched keypoints by our method

Figure 4. Remove mis-matched key points

Therefore, the proposed method performs better in distinguishing transformed images and similar images.

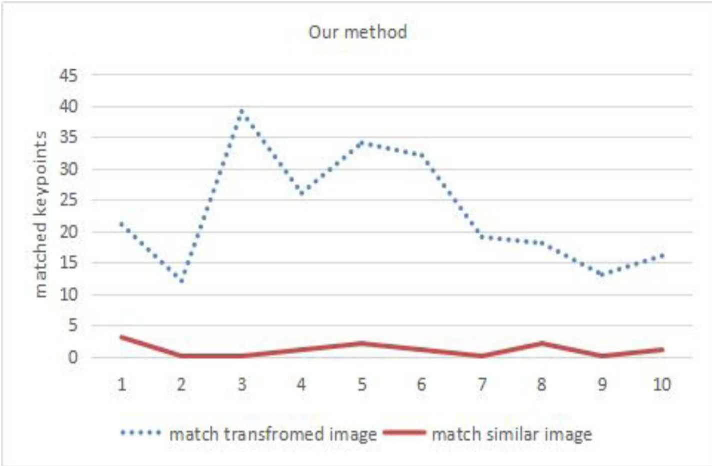
Image matching threshold of 10 is set up in this paper. If the number of correctly-matched key points between image I and partial image of query image is greater than or equal to 10, image I is a copy of query image Q.

3. Experimental results.

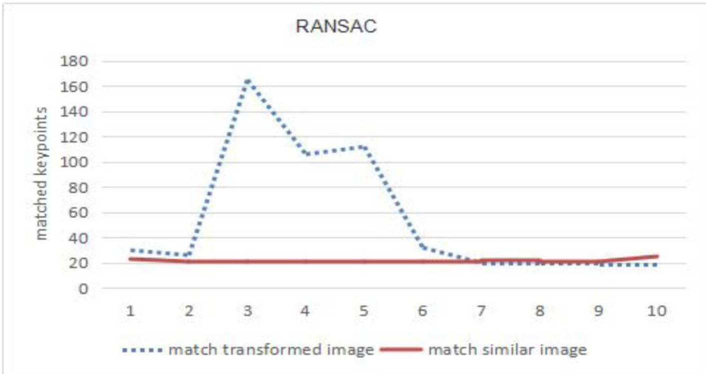
3.1. Image database. In this paper, an image database (IDB) containing 1000 images saved in JPEG format were taken from Baidu to serve as the test image database. The sizes of 1000 images for use were random, but all smaller than 800*800. In this paper, 15 transforms for each original image are adopted. The transforms are listed in Table 1.

3.2. Robustness of the proposed method. The thresholds used in following experiment were listed in Subsection 2.3. To prove the detection performance of the proposed method, the experiment is divided into two parts. In part 1, the whole original image is transformed by transforms in Table 1. In part 2, an important object is extracted from original image, and only this object is transformed. In Figure 6, a group of copies based on one original image for experiment part 1 and a group of copies based on one important object from another original image for experiment part 2 are listed.

The results are listed in Table 2. Because the proposed method segments the query image into a series of partial images, and each partial image contains one important object, a high detection ratio is obtained under picture in picture transform. Furthermore, compared to Kims method [1] and Lin et al.s method [6], the proposed method performs better under blurring, Gaussian noising, water coloring , mosaic tile and rotation. Under these transforms, the ratios of the proposed method are 1.0. It means every copy is detected. As shown in the last column of Table 2, in the second part of the experiment, though the copies only contains the main object of original image, all the copies are detected by the proposed method, because partial images of query image are matched with images in the proposed method.



(a) The number of matched keypoints with our method



(b) The number of matched keypoints with RANSAC

FIGURE 5. Distinction of our method and RANSAC

3.3. **Precision rate and recall rate.** In this section, 10 images were chosen, and each of them was transformed into 15 copies. Every 15 copies were regarded as one group. After insertion of the 15 groups of copies (150 new images), there were totally 1150 images in the database. The precision rates and recall rates are computed by Eq.(7) and Eq.(8).

$$Precision = \frac{number of copies detected whose N\_matched < thres\_match}{number of detections whose N\_matched < thres\_match} \tag{7}$$

$$Recall = \frac{number of copies detected whose N\_matched < thres\_match}{number of total copies} \tag{8}$$

where $N\_matched$ represents the number of matched key points. The threshold is 10, as shown in Subsection 2.3.

As shown in Figure 7, the precision rates of all the groups are 1, it means every returned image is a copy. On the other hand, the recall rates are all higher than 0.90.

4. **Conclusions.** A robust copy detection method on image is proposed in this paper. When detecting copied image, firstly, several salient regions in a query image are extracted based on the improved salient region detection. Secondly, surf key points in every partial image and in images in database are extracted and the number of matching key points

(a) A group of copies for experiment part 1

(b) A group of copies for experiment part 2

FIGURE 6. Two groups of copies: The top image in (a) is the original image, and the below 15 images are the copies generated from the original images. The upper left image in (b) is the original image and the upper right image is an important object extracted from original image. The 15 below images are the copies generated from the important object.

is computed based on the improved surf key points matching method. If the number of matching key points is greater than the threshold value, this image is considered a copy of query image. The experimental results show that the method proposed in this paper can detect copied images of a query image even when the image is cropped, rotated, put in picture, water-colored, changed in contrast, blurred or inserted with words. In the future, with the increase of the images in the database, we will use modern computing resources, such as GPU, to speed up the algorithm [5].

Table 2. Detection results compared with other methods

Experiments - part 2				Experiments - part 1
Transform No.	Our method	Lin et al.s method[6]	Kims method[1]	Our method
No transform	1.0	1.0	1.0	1.0
C1	1.0	-	-	1.0
C2	1.0	1.0	1.0	1.0
C3	1.0	-	-	1.0
C4	1.0	1.0	1.0	1.0
C5	1.0	1.0	1.0	1.0
C6	1.0	1.0	0.8	1.0
C7	1.0	0.6	1.0	1.0
C8	1.0	1.0	1.0	1.0
C9	1.0	1.0	1.0	1.0
C10	1.0	1.0	0	1.0
C11	1.0	1.0	1.0	1.0
C12	1.0	-	-	1.0
C13	1.0	0.7	1.0	1.0
C14	1.0	0.7	0.8	1.0
C15	1.0	0.7	1.0	1.0

centerNote: '- represents there is no such experimental results listed in [1] and [6]. center

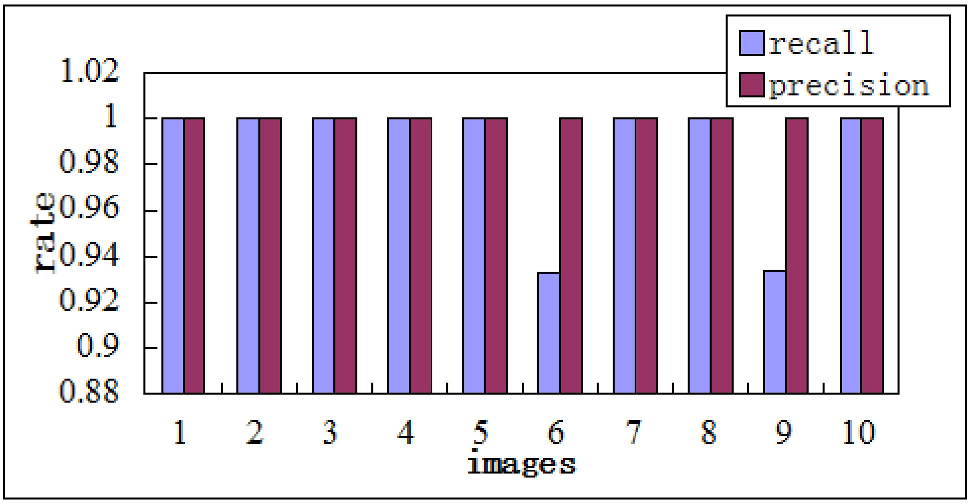


Figure 7. Recall rates and precision rates

Acknowledgment. This work was partially supported by National Natural Science Foundation of China (No.61370218).

REFERENCES

[1] C. Kim, Content-based image copy detection, *Signal Processing Image Communication*, vol.18, no.3 pp. 169C-184, 2003.

[2] M. N. Wu, C. C. Lin, C. C.Chang, Novel image copy detection with rotating tolerance, *Journal of Systems & Software*, vol.80, no.7, pp.1057–1069,2007.

[3] N. Baaziz, M. Guinin, Content-based image copy detection using dual signatures, *Proceedings of the 2011 IEEE International Symposium on Signal Processing and Information Technology IEEE Computer Society*, Bilbao, SPAIN, pp.017–022, 2011.

[4] N. Thanirat, S. Ngamsuriyaroj, The Effect of Attacks on DCT-Based Features for Image Copy Detection, *2014 IEEE 17th International Conference on Computational Science and Engineering (CSE)IEEE Computer Society*, ChengDu, China, pp.1483–1488, 2014.

[5] J. Baber, S. Satoh, C. Keatmanee, N. Afzulpurkar, Improving the performance of SIFT and CSLBP for image copy detection, *Telecommunications and Signal Processing (TSP), 2013 36th International Conference on IEEE*, Rome, Italy, pp.803–807, 2013.

[6] C. C. Lin, N. Klara, C. J. Hung, An image copy detection scheme based on edge features, *Multimedia and Expo, 2008 IEEE International Conference on IEEE*, Hannover, Germany, pp. 665–668, 2008.

[7] JL.Yao, XH.Wang,RB.Wang, Copy image detection based on local key points // Soft Computing and Pattern Recognition (SoCPaR), *011 International Conference of IEEE*, pp.256–262, 2011.

[8] M. A. Fischler, R. C. Bolles, Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography, *Communications of the ACM*, vol.24, no.6, pp. 381–395, 1981.

[9] M. Cheng, N. J. Mitra, X. Huang, P. H. S. Torr, S. Hu, Global Contrast Based Salient Region Detection, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.37, no.3, pp. 569–582, 2015.

[10] F. Perazzi, P. Kr?henbhl, Y. Pritch, A. Hornung, Saliency filters: Contrast based filtering for salient region detection, *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on IEEE*, Providence, RI, pp.733–740, 2012.

[11] Z. Wang, A. C. Bovik, H. R. Sheikh, E. P. Simoncelli, Image Quality Assessment: From Error Visibility to Structural Similarity, *IEEE Transactions on Image Processing*, vol.13, no.4, pp. 600–612, 2004.

[12] P. C. Niedfeldt; R. W. Beard, Convergence and Complexity Analysis of Recursive-RANSAC: A New Multiple Target Tracking Algorithm IEEE Transactions on Automatic Control,*A New Multiple Target Tracking Algorithm IEEE Transactions on Automatic Control*, vol.61,no.2,pp. 456–461, 2016.

[13] M. Alehdaghi, M. A. Esfahani, A.d Harati,Parallel RANSAC: Speeding up plane extraction in RGBD image sequences using GPU. *The 5th International Conference on Computer and Knowledge Engineering (ICCKE)*,2015.