

Very Fast Concentric Circle Partition-Based Replica Detection Method

Ik-Hwan Cho¹, A-Young Cho¹, Jun-Woo Lee¹, Ju-Kyung Jin¹, Won-Keun Yang¹,
Weon-Geun Oh², and Dong-Seok Jeong¹

¹ Dept. of Electronic Engineering, Inha University, 253 Yonghyun-Dong, Nam-Gu, Incheon,
Republic of Korea

² Electronics and Telecommunication Research Institute, 138 Gajeongno, Yuseong-Gu,
Daejeon, Republic of Korea

{teddydino, ayoung, jjunw6487, jukyung77, aida}@inhaian.net,
owg@etri.re.kr, dsjeong@inha.ac.kr

Abstract. Image replica detection becomes very active research field recently as the electronic device such as the digital camera which generates digital images spreads out rapidly. As huge amount of digital images leads to severe problems like copyright protection, the necessity of replica detection system gets more and more attention. In this paper, we propose a new fast image replica detector based on concentric circle partition method. The proposed algorithm partitions image into concentric circle with fixed angle from image center position outwards. From these partitioned regions, total of four features are extracted. They are average intensity distribution and its difference, symmetrical difference distribution and circular difference distribution in bit-string type. To evaluate the performance of the proposed method, pair-wise independence test and accuracy test are applied. We compare the duplicate detection performance of the proposed algorithm with that of the MPEG-7 visual descriptors. From experimental results, we can tell that the proposed method shows very high matching speed and high accuracy on the detection of replicas which go through many modification from the original. Because we use the hash code as the image signature, the matching process needs very short computation time. And the proposed method shows 97.6% accuracy on average under 1 part per million false positive rate.

Keywords: Image Replica detection, Image Retrieval, Concentric Circle Partition.

1 Introduction

Image retrieval (IR) system has been studied in various research fields and utilized in various real applications as digital images can be distributed easily through internet. Image retrieval system can be used in image indexing, searching and archiving. These enormous scientific interest leads to several international standards activity and the growth of related research areas. ISO/IEC MPEG standard group published MPEG-7 [1] standard and JPEG did JPSearch [2]. As a subset of image retrieval applications,

the interest for replica detection is gradually increasing. In general, all instances of reference image are called 'replicas'. And 'non-replicas' mean images which are not related to the modified versions of the reference image [3]. The reason why the replica detection is important and getting much attention in real application is that it can be the better alternative over the conventional contents protection algorithm such as encryption and watermarking. Cryptograph can encrypt digital contents with invisible user key and therefore only legal users who have key can decrypt contents appropriately. However it is very dependent on user key and encryption algorithm and both distributor and user need to equip particular tools for encryption and decryption. It could limit the convenient contents usage of user. On the other hand the watermarking method gives no limitation for the usage of contents, the watermark to be embedded in contents decreases the quality of original contents. Above all, digital watermarking technology is very weak when the watermarked image is modified or attacked by some distortions. On the contrary, image replica detection method has no limitation of user's contents usage as well as no distortion of original contents. Therefore it could resolve disadvantages of cryptograph and watermarking technologies.

There are some difference between image retrieval and replica detection. The objective of image retrieval system is to find out similar images according to particular features. For example, if we use the photo image of Miami Beach with blue ocean and white beach as a query image, the image retrieval system shows some beach images with blue colored ocean and white colored sand. The results may include images of pacific coast or Caribbean Sea. In contrast to the image retrieval, the image replica detection system should find out only the modified versions of input image exactly. While image retrieval system shows several results if some modified images are used as the query, replica detection system presents only their original images.

1.1 The Related Works

In order to detect original image for various replicas to be applied, key point-based method can be used [4]. Key point-based method detects feature points in image and describes local feature for each point. And it measures similarity by matching local descriptors of all feature point blocks. These methods show very good performance for replica detection and precise feature point detection methods such as SIFT [5] and Harris-Laplacian [6] can be proper solutions for these purpose. But they need high computational complexity for extracting and matching signatures. In real application, we need to process a large number of images so that extraction and matching speed are very important. In extraction and matching speed, matching speed is more important since signature extraction is generally carried out off-line. However matching process between query and reference images should be processed in real-time. Therefore we should consider speed problem in design of replica detection system. In addition, there have been many methods especially for replica detection system [7, 8, 9]. Alejandro et al. uses interesting point detection method to align query and reference images [7]. After global alignment between two images, block-based intensity difference is measured to obtain similarity score. Although it shows good performance under its own experiment condition, relative small size of database

can be the limitation in recent environment. Yan et al. proposed near-duplicate detection algorithm [8] and they also used local interesting point detection algorithm of Difference of Gaussian (DoG) detector [5]. And for interest point representation, PCA-SIFT is used as compact form relative to original one [10] and they use Locality Sensitive Hashing (LSH) to index local descriptor [11]. It shows excellent result for near duplicate detection, but it still has limitations of relative low matching speed. They improve matching speed dramatically rather than conventional method, it couldn't outperform simple Hamming distance measure. Recently Wang et al. proposed duplicate detection method using hash code [9]. Because of using very short binary hash code they can show very fast matching speed without even additional optimization method. However the proposed method in [9] uses too short hash code so that it can identify only small set of image pairs. As the size of database used in experiment is growing more and more, its performance decreases gradually.

1.2 The Proposed Replica Detection Model

As we commented in above section, replica detection system is a little different from image retrieval process. And there is also difference between conventional replica or duplicate detection method. In general replica detection algorithm has its own simulation model and they are usually separated into two groups (Fig. 1). Let's assume there are original images and their copied and transformed images. In most of conventional models, original images were used as query ones and its performance was measured by precision and recall. But in real application this model can be very exhaustive process because reference (transformed) image pool is usually extraordinary bigger than query (original) image pool so that matching process needs longer time for full searching. If we want to know how many replicas for specific original image is distributed, this model is reasonable. Therefore we can assume another replica detection model and the different replica detection model are proposed in this paper. the new replica detection model is to use transformed images as queries and original images as references. By using the proposed model we have fixed and relative small reference pool so that it takes short time to match query and reference. Even if there are so many images to be used as query, we can choose only image which we need to use as query. As a point of view of copyright protection, which is one of major application, we can collect reference images which want to be protected from copyright infringement and we choose specific suspicious image as query one. Then replica detection system searches its original from reference pool by matching query and all references. In this paper we employ the proposed replica detection model so that it is not proper to use conventional performance metrics such as precision and recall. Then we can define the modified requirements of replica detection system; robustness, independence and fast matching. For robustness, image signature must not be modified even if image is transformed or degraded. And for independence, if two images are perceptually different, the signatures of two images should be quite different. Finally for fast matching, matching process between two images should be carried out based on only signatures without image itself and matching algorithm must be considerably simple.

In this paper, we propose new fast replica detection algorithm which is based on concentric circle partition. We develop concentric circle partition method to make

rotation invariant signature. Based on concentric circle partition, we extract 4 kinds of features and they are all rotation and scaling invariant. In addition, the proposed method uses simple mathematical function to map scalar distribution into binary hash code. Binary hash code is very useful in matching process since it requires only XOR operation which needs smallest computation power. We define new performance measure to evaluate independence and robustness of the proposed method in experiment step. From these new performance metrics, we can identify the proposed method is very useful in practical application.

This paper is organized as follows. In section 2, we describe extraction process of the proposed signature in detail. And matching process is depicted shortly in section 3. Section 4 shows experiment results of the proposed method and we conclude it in section 5 and section 6.

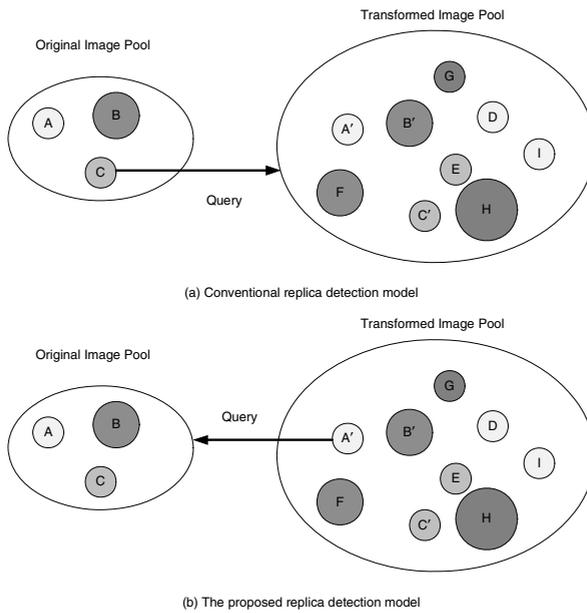


Fig. 1. Image replica detection model. (a) is conventional replica detection model which use original image as query and (b) is the proposed replica detection model which use original images as references. The proposed model requires short matching time relative to conventional model since it uses relative small and fixed reference pool.

2 Extraction of the Proposed Signature for Replica Detection

The proposed replica detector is based on concentric circle partition of input image. Fig. 2 shows example of concentric circle. Concentric circles are a set of circles which has common center and different radius. Prior to extracting feature from input image, concentric circle partition is applied to image and features are extracted from these segmented regions. Concentric circle is quantized into several sub-regions by

radius and angle level. In Fig. 2, left images shows basic concentric circle quantized by radius and right one is sub-regions quantized by different angles. The proposed method utilizes these sub-regions to extract features.

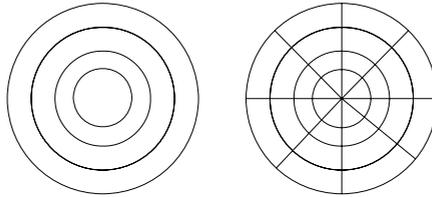


Fig. 2. Concentric circle and its partition in according to radius and angle

Basically, the proposed replica detector needs fixed number of bits and these hash type signature has advantages for signature size and matching speed. In addition, just one bit is allocated into one circle for signature. Actually overall replica detector has four kinds of signatures and they all are based on same concentric circle region. They are signatures of bit-string type and each signature has same simple hash generation method. Fig. 3 represents the overall block diagram for extracting features.

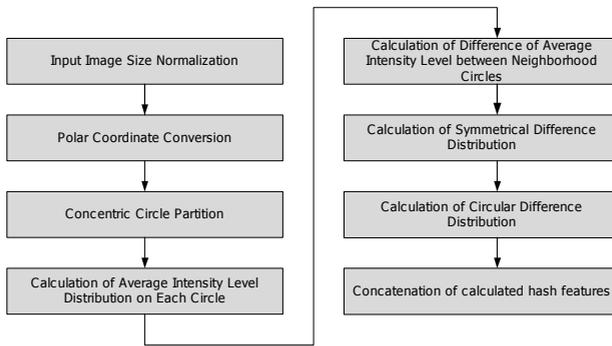


Fig. 3. The overall block diagram for signature extraction process

To make robust replica detector, we use four kinds of features based on concentric circle regions. The four features are average intensity level, difference of average intensity level, symmetrical difference and circular difference distribution. And they are represented by Hash type bit-string with same length. In following sections, the detailed extraction method of four features is explained.

2.1 Concentric Circle Partition

Input images are resized into fixed size while width-height ratio is remained. The fixed size is determined in according to diameter of largest circle and minimum value between resized width and height is equal to diameter. Resize operation is carried out by using bi-linear interpolation method. We used 256 as diameter in experiment.

To extract feature, concentric circle region concept should be applied into real image. Concentric circle is implemented by coordinate conversion from Cartesian coordinate to polar coordinate. Cartesian coordinate of (x, y) is converted into polar coordinate by using Eq. (1).

$$\begin{aligned}
 x &= r \cos \theta \\
 y &= r \sin \theta \\
 r &= \sqrt{x^2 + y^2} \\
 \theta &= \begin{cases} \arctan\left(\frac{y}{x}\right) & \text{if } x > 0 \text{ and } y \geq 0 \\ \arctan\left(\frac{y}{x}\right) + 2\pi & \text{if } x > 0 \text{ and } y < 0 \\ \arctan\left(\frac{y}{x}\right) + \pi & \text{if } x < 0 \\ \frac{\pi}{2} & \text{if } x = 0 \text{ and } y > 0 \\ \frac{3\pi}{2} & \text{if } x = 0 \text{ and } y < 0 \end{cases}
 \end{aligned} \tag{1}$$

After conversion into the polar coordinate, each pixel in (x, y) position has angle and distance from center and then polar coordinate map is obtained. The calculated polar coordinate map is quantized by angle and radius. In this experiment, we use 32 radius level and 36 angle level.

2.2 Average Intensity Level Distribution

First feature of the proposed signature is average intensity level in each circle. For all circle regions, average intensity level for each circle is calculated and the distribution from inner circle to outer circle is obtained (Eq. (2)).

$$C_i = \frac{1}{angleLevel} \sum_{j=0}^{angleLevel} P_{i,j} \tag{2}$$

$P_{i,j}$ is all pixel value in i th radius level and j th angle level.

2.3 Difference Distribution of Average Intensity Level

Second feature is difference distribution of average intensity level distribution which is calculated in previous step. Its mathematical representation is depicted in Eq. (3).

$$V_i = abs(C_{i+1} - C_i) \tag{3}$$

C_i is all average intensity level in i th radius level.

2.4 Symmetrical Difference Distribution

Third feature is symmetrical difference distribution. Symmetrical difference is calculated by summing absolute differences between average level for some angle

region and average value of other side region in each circle. In Fig. 4(a), two gray regions are symmetrical regions each other in one circle and so symmetrical difference for each regions is obtained by calculating absolute difference between average intensity value of two same symmetrical regions. Its mathematical representation is shown in Eq. (4).

$$S_i = \frac{1}{angleLevel/2} \sum_{j=0}^{angleLevel/2-1} abs(C_{i,j} - C_{i,j+angleLevel/2})$$

$C_{i,j}$ is all average intensity level in i th radius level and j th angle level.

2.5 Circular Difference Distribution

Final fourth feature is circular difference distribution. Circular difference in one circle is calculated by summing absolute difference between average intensity level in some angle region and average intensity value in counter-clockwise directional neighborhood angle region. In Fig. 4(b), gray regions are neighborhood regions in one circle and their absolute difference is circular difference for two regions. Mathematical representation of circular difference is shown in Eq. (5).

$$R_i = \frac{1}{angleLevel} \sum_{j=0}^{angleLevel-1} abs(C_{i,j} - C_{i,(j+1) \bmod angleLevel})$$

$C_{i,j}$ is all average intensity level in i th radius level and j th angle level.

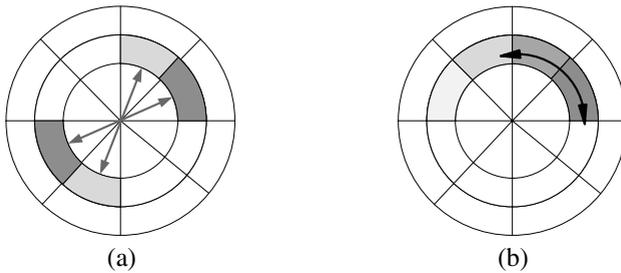


Fig. 4. (a) Symmetrical regions in one circle, circular difference calculation for counter-clockwise

2.6 Merging of Hash Type Features

In previous signature extraction process, we obtained four kinds of distribution and its distribution is spread from inside to outside of concentric circles. In this step, simple

hash table is applied to this distribution to change scalar distribution into binary bit-string. Eq. (6) is simple Hash mapping function used to represent graph pattern with binary string.

$$B_i = \begin{cases} 1, & M_{i+1} > M_i \\ 0, & M_{i+1} \leq M_i \end{cases} \quad (6)$$

M_i is distribution value in i th index.

This mapping function is applied into pre-calculated distribution values in each signature independently. And extracted bit-string is final proposed replica detector. In addition, above all process can be repeated with different parameter and they are mixed as final replica detector.

3 Matching Process

Matching of the proposed signature is carried out by measuring general normalized Hamming distance simply. Matching process of the proposed signature is very simple since its signature has just bit-string, not real value to be required of Euclidean distance calculation or additional processing. Eq. (7) shows distance measure used in the proposed method.

$$D = \frac{1}{N} \sum_{j=0}^{N-1} R_j \otimes Q_j \quad (7)$$

N is the number of bits of descriptor

and R_j and Q_j is bit of reference and query in j th index.

4 Experimental Results

For performance evaluation of the proposed replica detector, two kinds of experiment are carried out. First one is simulation to get distance threshold value under 1 ppm (parts per million) false positive error constrictioin. And second one is experiment to measure accuracy of the proposed replica detector. Accuracy is evaluated by using only Correct Retrieval (CR) Ratio. Since we cannot use traditional precision and recall, we define measurement metrics. To define CR ratio, let us assume that there are M query images for one ‘transformed’ version (e.g. blur). To compute CR ratio, the number of true pairs that are classified as containing copies (i.e. the original image and its modified version) is counted (K). The CR ratio is defined as:

$$\text{Correct Retrieval Ratio} = \frac{K}{M}$$

The query shall be performed for all original images and the experiment repeated for all types of modifications. For this experiment, 60,555 images are used for

experiment independence test. This larger database has various natural photo images including several genre including animal, people, food, architecture, house and abstract image [12]. And 3,943 original images and 23 kinds of modifications are used. Modification list is depicted in Table 1. Actually feature extraction process is repeated twice with different radius levels and two extracted bit-strings are mixed. In this experiment for the proposed replica detector, we set up basic parameters as following;

- Circle diameter = 256
- Circle radius level = 32, 16
- Circle angle level = 36

For comparison of performance, we implement duplicate detection algorithm based on PCA-HAS in [9] and two MPEG-7 visual descriptors of color layout and edge histogram descriptor as references. We use same database for these three reference algorithms and carry out same independent and accuracy test. And all experiment is processed in 3.4GHz Intel Pentium 4 Processor with Microsoft Windows XP.

4.1 Independent Test

When cross similarity measurement with 60,555 images is processed, threshold distance under 1ppm false positive rate is 1,833th distance. Fig. 5 shows cross distance histogram of all methods including the proposed one for very large database.

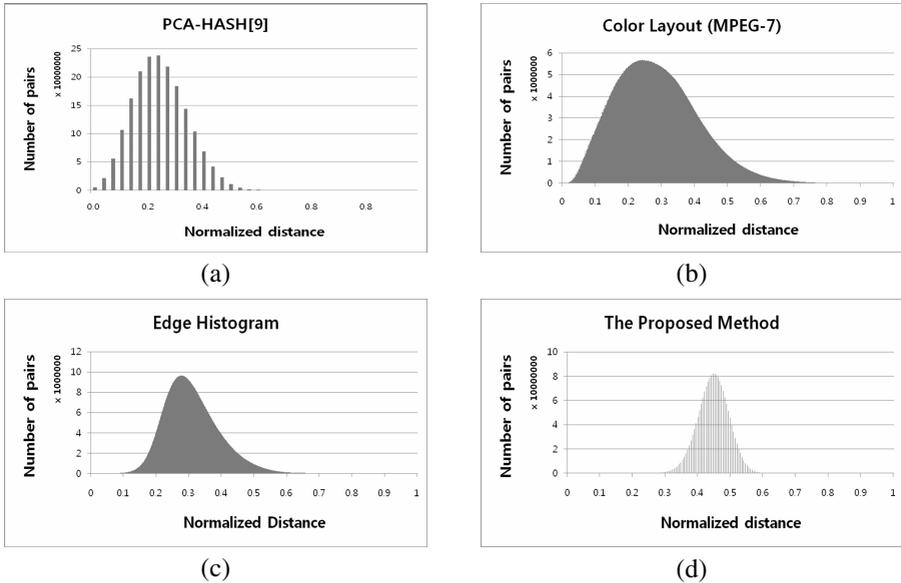


Fig. 5. Cross distance histogram of reference methods and the proposed method. Threshold distance of each algorithm is as following; ((a): 0.0315, (b):0.007, (c):0.036, (d):0.141).

4.2 Accuracy

Accuracy is measured by average CR Ratio and results under 1 ppm false positive rate are represented in Table 1. In experiment results, the proposed method shows very high accuracy of average 97.6 %.

Table 1. CR ratio results for modification levels

		PCA-HASH [9]	Color Layout (MPEG-7)	Edge Histogram (MPEG-7)	The Proposed Method
1	Add Gaussian Noise ($\sigma=2.5$)	0.530	0.678	0.481	0.993
2	Add Gaussian Noise ($\sigma=4.5$)	0.525	0.625	0.238	0.983
3	Add Gaussian Noise ($\sigma=8.0$)	0.517	0.532	0.075	0.971
4	Auto level	0.382	0.055	0.940	0.992
5	Blur (3x3)	0.525	0.738	0.091	0.998
6	Blur (5x5)	0.524	0.734	0.052	0.995
7	Bright change (5%)	0.511	0.032	0.631	0.994
8	Bright change (10%)	0.478	0.002	0.356	0.989
9	Bright change (20%)	0.396	0.000	0.124	0.967
10	JPG compress (qf=95%)	0.537	1.000	0.998	1.000
11	JPG compress (qf=80%)	0.534	0.865	0.390	0.998
12	JPG compress (qf=65%)	0.526	0.722	0.203	0.991
13	Color reduction (8bit)	0.493	0.224	0.256	0.983
14	Color reduction(16bit)	0.466	0.223	0.445	0.980
15	Histogram equalization	0.296	0.063	0.213	0.919
16	Monochrome	0.530	0.144	0.273	0.997
17	Rotation (10°)	0.038	0.010	0.013	0.992
18	Rotation (25°)	0.004	0.002	0.013	0.847
19	Rotation (45°)	0.001	0.001	0.010	0.876
20	Flip	0.003	0.015	0.002	1.000
21	Rotation (90°)	0.000	0.001	0.000	0.997
22	Rotation (180°)	0.001	0.001	0.001	0.997
23	Rotation (270°)	0.000	0.001	0.000	0.999
Average		0.340	0.290	0.252	0.976

The above test shows that the proposed method is very robust for various transforms including rotation by arbitrary angle and flip. However the other method is less robust for image distortion and they show very low performance on rotation and flip. Fig. 6 shows the relation of their performance graphically.

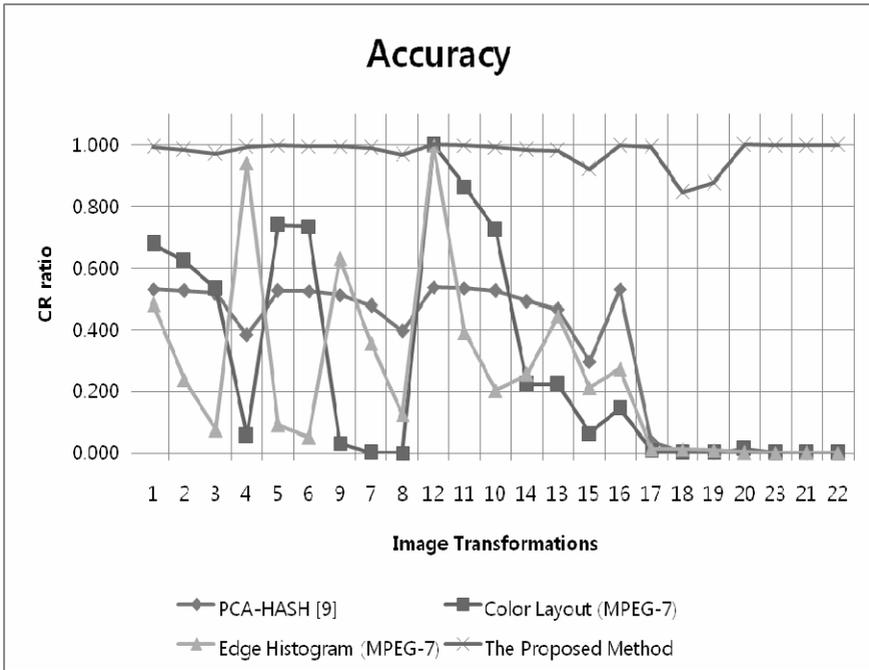


Fig. 6. The performance of all methods for variable distortions

4.3 Signature Size

The proposed replica detector is composition of four signatures and its length is related with radius level. In this experiment, we use 32 and 16 as radius level. As a result, total signature size is only 192 bits. As point a view of signature size, method proposed in [9] has smallest of only 32 bits. Table 2 shows size of the signature of all methods.

Table 2. Signature size (bits)

	Number of bits
PCA-HASH [9]	32
Color Layout (MPEG-7)	42
Edge Histogram (MPEG-7)	240
The Proposed Method	192

4.4 Computational Complexity

In this paper, we limit the concept of computational complexity into complexity in matching process. Because extraction process can be carried out in off-line in general, it requires no real time processing. In matching step, computation complexity is very related to matching method of the signature. PCA-HASH and the proposed method in [9] use binary string as signature so that they need only XOR operation. This XOR

operation is basic operation which needs minimum computing power in general computer architecture. Because of using binary hash code, they can obtain high matching speed relatively while the other methods use more complex matching scheme. Table 3 shows computational complexity of all methods as the number of matching pairs per second.

Table 3. Computation matching complexity

	The number of matching pairs per second
PCA-HASH [9]	90,000,000
Color Layout (MPEG-7)	4,230,543
Edge Histogram (MPEG-7)	293,206
The Proposed Method	15,000,000

5 Discussion

In this paper, we proposed new fast concentric circle partition-based signature method for image replica detection. The proposed method partitions image into concentric circle shape in according to radius and angle. Basically the final signature has 1 bit per each circle after partition. For each circle, four kinds of features are extracted and its values are converted into 1 bit by using simple hash function with the relation to neighborhood circle. Each four feature has different characteristics relatively. First feature of average intensity distribution represents overall pixel level distribution from center to outer region. This feature can show general characteristics of the image. However its diversity of intensity distribution in image has limitation, which means that it is possible for two different images to have similar distribution when the signature is presented as bit-string by using only magnitude difference between neighborhood circles. So we add difference of intensity distribution as second feature. This feature compensates the weakness of first feature by considering the degree of magnitude difference. Therefore second feature acts valuable role to increase discrimination between different images. Third and fourth features represent local characteristics while first feature describes image globally. Third feature describes the variation on symmetry through circles. And fourth feature describes the degree of variation to the circular direction in each circle.

By utilizing these four kinds of features together, the proposed method shows very high CR (Correct Retrieval) ratio of 97.6 % under 1 ppm false positive rate. In addition, the proposed signature has very small size of fixed 192 bits per image and shows very fast matching process 15 million pairs per second. In our experiment part, we compare the proposed method with other three algorithms. First one is method proposed in [9] for large scale replica-pair detection system. Its algorithm calculate block-based pixel mean as a vector and this vector is optimized and compressed into only 32 dimensions by Principle Component Analysis (PCA). Finally vector of 32 dimensions are converted into 32 bits hash code. However in our experiment it doesn't show good performance because it has too short signature. Only 32 bits is not enough to have unique and independent information of specific image. The MPEG-7 visual descriptors of color layout and edge histogram are selected as targets for

experimental comparison since they show better performance in every MPEG-7 visual descriptors relatively. As the experimental results shows, however, MPEG-7 visual descriptors which are developed for image retrieval applications cannot satisfy the requirement of replica detection. In case of edge histogram descriptor, it partitions image into fixed number of blocks so that they cannot synchronize signatures between query and reference if original image is rotated or flip geometrically. Color layout descriptor has critical limitations even though it shows good performance for some image distortions. It uses color information in overall range of image so that it cannot protect original information when it loses color information by converting gray scale image. Therefore it is difficult to expect high performance of color descriptor for variable image modifications. For aspect of computation complexity, the proposed method outperforms these two descriptors since they use Euclidean distance measure and it requires more computational power rather than Hamming distance measure using basic XOR operation.

For the development of better image replica detector, most important factors are high accuracy and low computational complexity. Digital image is usually copied and modified without large efforts through simple software. Therefore the replica detector should cover variable modification of source image. In this paper, we apply 23 kinds of transformations and they can be considered as the most frequent usage by normal users in real application. For these various modifications, the proposed method shows very high performance regularly for all modifications. And In general, best method which shows the fastest speed in matching of feature-based signature is to describe feature with bit-string and use simple Hamming distance in matching procedure. In computational calculation, it needs just one clock to carry out XOR operation which other operation needs more clocks. However bit-string type signature has low accuracy relative to other type method since its bit-string can show just binary information. Binary information represents only two types of information; ON or OFF. However these problems of bit-string type signature can be compensated by using proper design of Hash table. In this paper, we firstly describe feature with floating values and obtain final signature by presenting its distribution into bit-string. The proposed method shows high performance for accuracy and speed by mixing various features appropriately.

6 Conclusion

We proposed very fast replica detection method based on concentric circle partition of image. The proposed method utilizes the partitioned regions and extracts four kinds of features which have different characteristics to describe image efficiently. The proposed method shows very high CR (Correct Retrieval) ratio of 97.6 % under 1 ppm false positive rate. And the mixed signature has the form of bit-string by applying simple Hash table which leads to fast matching speed. The main advantages of the proposed replica detector are high accuracy and matching speed. Therefore the proposed method will be very useful in real application such as the detection of illegal image replicas.

Acknowledgments. The presented research is supported by INHA UNIVERSITY Research Grant.

References

1. Martinez, J.M., Koenen, R., Pereira, F.: MPEG-The Generic Multimedia Content Description Standard. *IEEE Multimedia* 9(2), 78–87 (2002)
2. ISO/IEC/JTC1/SC29/WG1 WG1 N3684, JPSearch Part1 TR - System Framework and Components (2005)
3. Maret, Y., Dufaux, F., Ebrahimi, T.: Image Replica Detection based on Support Vector Classifier. In: *Proc. SPIE Applications of Digital Image Processing XXVIII*, Santa Barbara, USA (2005)
4. Ke, Y., Sukthankar, R., Huston, L.: An Efficient Parts-Based Near-Duplicate and Sub-Image Retrieval System. In: *ACM International Conference on Multimedia*, pp. 869–876 (2004)
5. Lowe, D.G.: Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision* 60, 91–110 (2000)
6. Mikolajczyk, K., Schmid, C.: Scale & Affine Invariant Interest Point Detectors. *International Journal of Computer Vision* 60, 63–86 (2004)
7. Jaimes, A., Shih-Fu, C., Loui, A.C.: Detection of non-identical duplicate consumer photographs, *Information, Communications and Signal Processing, 2003 and the Fourth Pacific Rim Conference on Multimedia*. In: *Proceedings of the 2003 Joint Conference of the Fourth International Conference*, vol. 1, pp. 16–20 (2003)
8. Ke, Y., Sukthankar, R., Huston, L.: Efficient Near-duplicate Detection and Sub-image Retrieval. In: *Proc. ACM Intl. Conf. on Multimedia*, New York, pp. 869–876 (2004)
9. Wang, B., Li, Z., Li, M., Ma, W.Y.: Large-Scale Duplicate Detection for Web Image Search. In: *IEEE International Conference on Multimedia and Expo.*, pp. 353–356 (2006)
10. Ke, Y., Sukthankar, R.: PCA-SIFT: A More Distinctive Representation for Local Image Descriptors. In: *Proceedings of IEEE Computer Vision and Pattern Recognition* (2004)
11. Indyk, P., Motwani, R.: Approximate Nearest Neighbor-towards Removing The Curse of Dimensionality. In: *Proceedings of Symposium on Theory of Computing* (1998)
12. Mammoth – 800,000 Clipart DVD by FastTrak, <http://www.amazon.co.uk>