THE INSTITUTE OF ELECTRONICS,
INFORMATION AND COMMUNICATION ENGINEERS

TECHNICAL REPORT OF IEICE.

Copyright Protection of Images Based on Large-Scale Image Recognition

Koichi KISE†, Satoshi YOKOTA†, and Akira SHIOZAKI†

† Department of Computer Science and Intelligent Systems,
Graduate School of Engineering, Osaka Prefecture University
E-mail: †{kise,shiozaki}@cs.osakafu-u.ac.jp, ††yokota@m.cs.osakafu-u.ac.jp

Abstract This paper presents a method of copyright protection for digital still images based on their recognition with a large-scale image database. For the purpose of copyright protection of images, digital watermarking has been extensively studied for more than a decade. However, it is still difficult to achieve enough robustness against alternations such as StirMark attacks. In order to solve this problem, we propose a method of copyright protection based on new image recognition technologies that are tolerant to various image distortions including geometric transformations. The image recognition enables us to find a corresponding image from at least tens of thousands of reference images to be protected in response to a query image that undergoes StirMark attacks and digital watermarking. From experimental results using 10,000 images in the database, we have confirmed that the method is capable of recognizing 99.5% of query images only within 216 ms / query (excluding feature extraction).

Key words copyright protection, image recognition, StirMark, digital watermarking, SIFT

1. Introduction

Copyright protection of digital images are one of the most important problems to be solved to promote the sound development of the Internet. Digital watermarking (simply called watermarking, in the following) has been regarded as a promising way for fulfilling this requirement [1].

Watermarking has attracted researchers and users because it offers convincing scenarios for copyright protection such as follows. With the help of watermarking, one can embed information about copyright, and a purchaser, etc. to the image to be protected. When it is illegally distributed, the copyright holder can extract embedded information from a distributed image. This discourages pirates to make illegal copies.

In order to bring such a scenario into reality, some properties are required to watermarking methods. First, the watermark should be imperceptible to keep the image valuable. Second, the watermark cannot be removed from the image unless the removal process also debases the image.

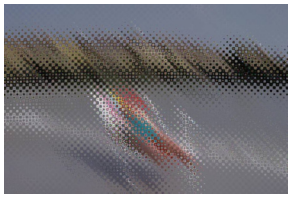
For most of currently available methods, however, these properties are partially held. An important limitation is that many methods cannot survive geometric transformations including cropping, scaling, rotation and shear unless the user gives up the first property.

In this paper, we propose a totally different solution for the copyright protection based on state-of-the-art technologies of image recognition with local descriptors such as SIFT (Scale-Invariant Feature Transform) [2] and PCA-SIFT (Principal Component Analysis SIFT) [3]. A local descriptor is a feature vector that describes a local region of the image. In general, local descriptors are invariant to geometric transformation such as scaling and rotation. In addition they are hardly affected by cropping due to their locality. Local descriptors extracted from an image in question are utilized by image recognition to determine whether or not it is copyrighted using a database of copyrighted images. The size of the database poses few problems for current recognition technologies. For example, it is reported in the literature that recognition of an unknown image by matching with a database of 100,000 images takes only 1ms [4], [5]. Once the image in question is known to be a copyrighted image in the database, we can employ its original image in the database to recover the embedded watermark.

We conducted preliminary experiments about recognition of watermarked and distorted images. Various distortions were generated by a standard benchmark program called StirMark [6]. We employed 10,000 images as watermarked and distorted images, and also 10,000 images for a database of original images. The results show that the recognition is robust (99.5% accuracy) and fast (216ms / image in question). Thus we have confirmed that the proposed method is a promising way to solve the problem of watermarking.

Figure 1 Original image.

(a) Cropping



(b) Fraction of pixel displacement

(c) Punch/Pinch

(d) Corner shift

Figure 2 Some image transformations implemented in StirMark.

2. Digital Watermarking and StirMark

Methods of watermarking can be divided into three types: blind, semi-blind, non-blind[1]. Blind watermarking means that neither original images with nor without the watermark are needed for detecting the watermark; only a publicly available watermarked image that may be distorted by image processing is needed. Semi-blind watermark indicates a way of watermarking that requires the original (clean) watermarked image. Non-blind watermarking requires the original image without the watermark.

A majority of research efforts have been paid to the development of blind watermarking, because it is easy to use; only the image in question is needed to detect watermarking. However, methods of this scheme are most difficult to develop and often insufficient to attacks such as StirMark.

StirMark is a standard way to benchmark watermarking methods. Basically it applies several image distortions including addition of noise, frequency filtering, global and local geometric transformations, and lossy (JPEG) compression. Some global and local geometric transformations applied to the original image in Fig. 1 are shown in Fig. 2[1].

When StirMark is in real use, such distortions are imperceptible to human, but severely damage watermarks by most

[1]The strength of these transformations is far beyond imperceptible simply for the illustrative purpose.

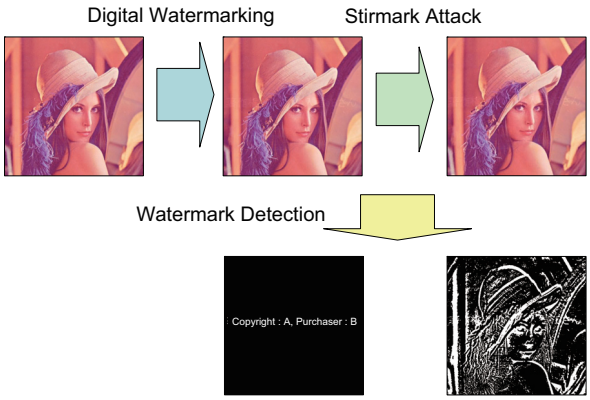


Figure 3 Digital watermarking and StirMark.

methods. An example is shown in Fig. 3. After a StirMark attack, the detection failed to find the information on the copyright holder and the purchaser of the copy. Thus an intermediate goal of the research on blind watermarking has been to develop a method that survives StirMark.

As compared to the blind scheme, semi-blind and non-blind schemes were pursued at early years of research on watermarking but are not popular in recent years. One of reasons for this would be that it requires matching between an image in question and each original copyrighted image. If the number of copyrighted images is large, the computational cost of matching is considered to exceed the limit of what can be paid in real applications. In other words, researchers would consider that these schemes do not scale well.

3. Copyright Protection Based on Image Recognition

3.1 Proposed Scheme

The purpose of this paper is to revive the non-blind scheme with the help of up-to-date technologies of image recognition that are highly scalable.

Figure 4 illustrates the overview of the proposed scheme that consists of three steps. At the step 1, an image in question, called a query image in the following, is recognized to find the corresponding copyrighted original image in the database. For the recognition, we utilize local descriptors each of which represents an invariant feature of a local region around a feature point. Note that feature points are determined based only on the image, basically insensitive to geometric transformations, and generally distributed over the image. If the original image is found, correspondence between feature points are obtained as a side effect of recognition as shown in Fig. 4. The next step is to normalize a query image based on the correspondence of feature points. This normalization process is to remove geometrical and other image distortions by comparing it to the original. The final step is to detect the embedded watermark from the normalized

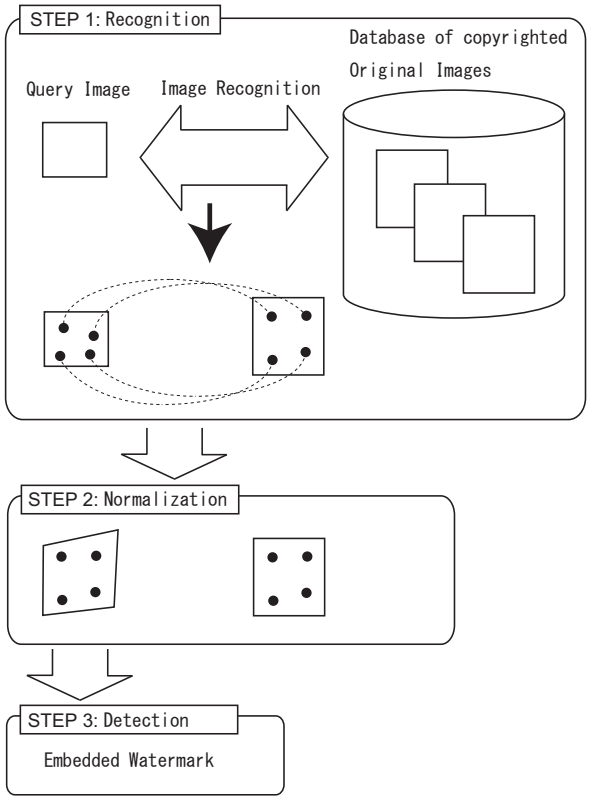


Figure 4 Proposed scheme.

query image.

As the first step of realizing this scheme, we focus on the step 1 in this paper.

3.2 Image Recognition

Image recognition has a long history of coping with image distortions including addition of noise, changes of pixel values, geometric transformations, and cropping (occlusion). Local descriptors developed in recent years are capable of solving such problems. In this paper, we employ PCA-SIFT [3], a variant of the well-known SIFT operator [2], which produces feature vectors of 36 dimensions.

A drawback of using local descriptors was a large computational load of matching them, i.e., computation of the distance between vectors to find nearest neighbors. It is typical to have about 2,000 descriptors from a single VGA image, the database of 10,000 images contains 20M descriptors. Thus the brute-force matching requires 40G ($= 2{,}000 \times 20\text{M}$) times of distance computation, which could need more than an hour for recognizing one image with a modern computer. However, recent technologies of finding nearest neighbors efficiently based on a tree and a hash structure, have made a breakthrough.

The proposed recognition utilizes an well-known method of nearest neighbor search called Approximate Nearest Neighbor, ANN [7]. ANN is capable of finding the nearest neighbor vector in an approximate way that enables us to cut the computation time drastically.

Recognition is realized based on the voting. Each feature vector extracted from the query image, or each query feature vector, has one vote. For each query feature vector, the recognizer casts a vote to the image whose feature vector is the nearest to the query feature vector. The original image which accumulates the maximum votes is the recognition result if some simple conditions such as the minimum number of votes fulfill predetermined conditions.

4. Experimental Results

4.1 Overview

We evaluated the robustness of the proposed image recognition against StirMark and embedding of digital watermarking, both of which disturb the recognition.

As the StirMark, we employed StirMark3.1. In addition to the default setting to generate imperceptible changes of the original image, we also applied several geometric distortions that cause perceptible changes for evaluating the robustness of recognition.

As watermarking methods, we applied SteganoEngine Ver.2.3, and Zhu's method [8] whose implementation is available at [9]. Zhu's method is based on the multiresolution obtained by the wavelet transformation. A watermarked image is given by

$$v_i' = v_i(1 + \alpha x_i) \tag{1}$$

where v_i is a high-pass wavelet coefficient, x_i is an element of the embedded watermark, and v_i' is a coefficient after the embedding. The strength of watermarking can be controlled by the parameter α.

By combining settings of StirMark and watermarking, we conducted the following three experiments.

Experiment 1 Robustness against imperceptible changes of images by the default StirMark

Experiment 2 Robustness against perceptible changes of images by several geometric transformations available at StirMark

Experiment 3 Robustness against watermarking by changing the strength α

As images in the database, we utilized 10,000 images which are as a part of those available at [10]. From each image, about 2,000 feature points were extracted. In order to generate query images, we randomly selected 100 images from the database, and applied watermarking and StirMark. In the following the selected 100 images are called the base images.

4.2 Experiment 1: Default StirMark

First, we evaluated the proposed recognition based on "normal" settings of watermarking and StirMark. We utilized SteganoEngine to embed information on the base images.

Table 1 Results of the experiment 1.

watermarking	without	with
accuracy	100%	99.49%
processing time	217ms	216ms

Then we applied StirMark to each base image in 100 different ways. Thus the number of generated query images is $100 \times 100 = 10,000$. Query images generated only by StirMark (without watermarking) were also employed for comparison. The results were evaluated using both the accuracy of recognition and the processing time excluding the time for feature extraction (PCA-SIFT).

Table 1 shows the results. Without the watermarking, the accuracy of recognition was 100%. This means that all 10,000 images were correctly associated with its original from the database of 10,000 images even under StirMark attacks. The average processing time was 217ms. For the images with watermarking, the accuracy dropped 0.51% but was still over 99%. There was no significant difference of the processing time.

From the results shown above, it is confirmed that the recognition method is robust against imperceptible image transformations such as StirMark and watermarking as well as fast enough as compared to the processing time for other watermarking methods.

4.3 Experiment 2: Various Geometric Transformations

One of the important problems for watermarking is how to cope with geometric transformations. It is not easy for most watermarking methods to achieve tolerance to severe geometric transformations. If the recognition method has the tolerance, it may be possible to make a watermarking method more robust to such transformations based on the result of the recognition.

The experiment 2 was designed to evaluate this aspect of the proposed method. For each base image, we applied 4 geometric transformations: "cropping", "fraction of pixel displacement", "punch/pinch", and "corner shift" shown in Fig. 2. As the level of transformations we did not keep them imperceptible but also made them perceptible for the purpose of testing the robustness of the recognition method. For each base image, we applied 43 different transformations, and obtained 4,300 query images in total.

The results are shown in Figs. 5–8 where the lines with squares and triangles indicate the accuracy for queries with and without watermarking, respectively. As clearly shown in these figures, recognition rates drop as the level of transformations become more perceptible. For the original image shown in Fig. 9, query images with watermarking are illustrated in Figs. 10 – 13. Except for the image in Fig. 10, not

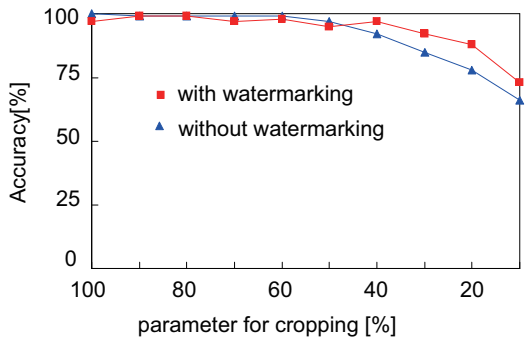


Figure 5 Results of experiment 2 : (a) Cropping .

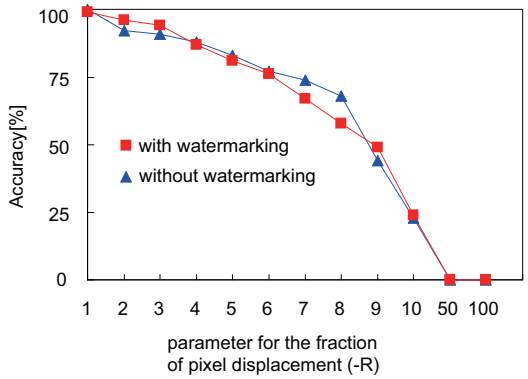


Figure 6 Results of experiment 2: (b) Fraction of pixel displacement.

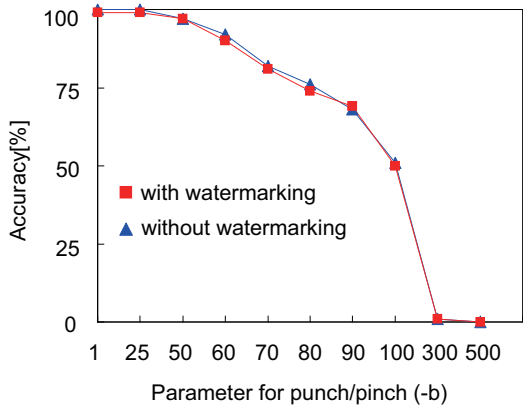


Figure 7 Results of experiment 2: (c) Punch/Pinch.

only successfully recognized images (a), but also failure images (b) are shown. The successful images are of the largest fluctuations under the condition that the images were successfully recognized. The failure images are, on the other hand, of the smallest fluctuations that yield misrecognition.

Most failure query images (b) indicate that their quality is extremely low as compared to the original, and thus need not to protect their copyrights. This can also be said for some successful images (a). From these results, we have confirmed that the recognition method is robust enough to various types of strong geometric transformations and thus dependable when we consider copyright protection based on the recognition.

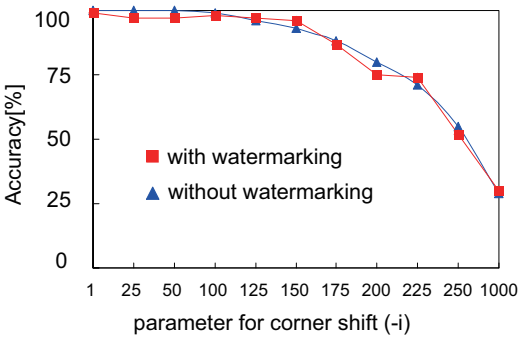


Figure 8 Results of experiment 2: (d) Corner Shift.

Figure 9 Original image.

(a) successful case : with the parameter `-i` 10.

Figure 10 A result of the cropping transformation.

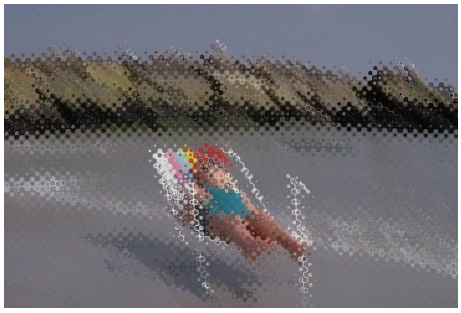
4. 4 Experiment 3: Watermarking

The last experiment was to test the tolerance to watermarking by varying the parameter α of Eq. (1).

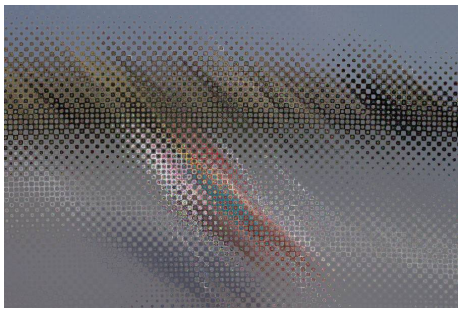
For this experiment, query images were prepared as follows. First, we embedded watermarks for each base image in ten different ways with parameters from 0.1 – 1.0. Thus the number of images with watermarks is $100 \times 10 = 1,000$. Next, we applied StirMark with the default parameters to each image with a watermark. The result is shown in Fig. 14. From $\alpha = 0.1$–0.3, 100% accuracy were obtained. An example of query images with $\alpha = 0.3$ is shown in Fig. 15(b). With $\alpha = 1.0$, the accuracy fell down to 90%. An example query is shown in Fig. 15(c). Because the quality was so deteriorated with larger values of α, it is not meaningful to consider such images for protection.

5. Conclusion

We have proposed a new scheme of copyright protection of images based on image recognition. The up-to-date technologies of recognition such as local descriptors and fast nearest neighbor search allow us to revive non-blind watermarking without losing its scalability. From the experimental results on large-scale image recognition of watermarked and distorted images, we have confirmed that there exist few problems of recognizing images distorted by default StirMark in both senses of accuracy and speed. It is also exemplified that



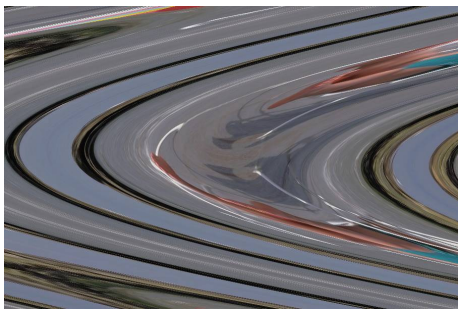
(a) successful case : with the parameter `-R` 10.



(b) erroneous case : with the parameter `-R` 50.

Figure 11 Results of the fraction of pixel displacement.

(a) successful case : with the parameter `-b` 100.



(b) erroneous case : with the parameter `-b` 300.

Figure 12 Results of punch/pinch.

(a) successful case : with the parameter `-i` 225.

(b) failure case : with the parameter `-i` 250.

Figure 13 Corner shift.

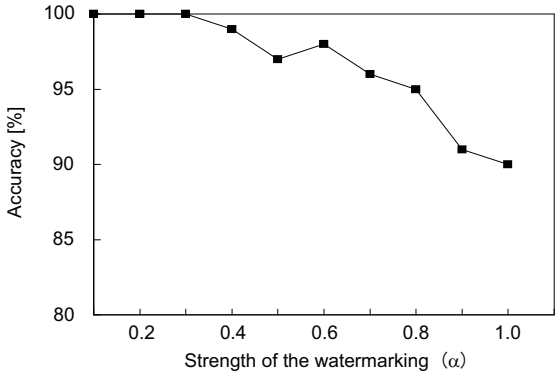


Figure 14 Results of the experiment 3.

(a) original

(b) $\alpha = 0.3$

(C) $\alpha = 1.0$

Figure 15 Query images for the experiment 3.

the recognition method survives severe distortions caused by StirMark and watermarking.

The future work includes the completion of the overall method to the detection of watermark in order to confirm the effectiveness of the proposed scheme.

Acknowledgement

This research was supported in part by the Grant-in-Aid for Scientific Research (B) (19300062) from Japan Society for Promotion of Science, Research for Promoting Technological Seeds from Japan Science and Technology Agency.

References

[1] http://www.watermarkingworld.org/.

[2] D. Lowe, "Distinctive image features from scale-invariant keypoints," International Journal of Computer Vision, vol.60, no.2, pp.91–110, 2004.

[3] Y.Ke and R.Sukthankar, A more distinctive representation for local image discriptors, IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Vol. 2, pp.506–513, 2004.

[4] K. Noguchi, K. Kise and M. Iwamura, Cascading approximate nearest neighbor searchers and its effects on object recognition, IEICE Technical Report, PRMU2007-44, pp.99–104, 2007. [In Japanese].

[5] K. Kise, K. Noguchi and M. Iwamura, Simple representation and approximate search of feature vectors for large-scale object recognition, Proc. of BMVC2007, 2007. [to appear].

[6] http://www.petitcolas.net/fabien/watermarking/stirmark/.

[7] S. Arya, D. M. Mount, N. S. Netanyahu, R. Silverman and A. Y. Wu, "An optimal algorithm for approximate nearest neighbor searching fixed dimensions," Journal of ACM, vol.45, no.6, pp.891–923, 1998.

[8] W. Zhu, Z. Xiong and Y. Q. Zhang, "Multiresolution watermarking for images and video," IEEE Transactions on Circuits and Systems for Video Technology, vol.9, no.4, pp.545–550, 1999.

[9] http://www.cosy.sbg.ac.at/~pmeerw/watermarking/.

[10] http://www.pittsburgh.intel-research.net/projects/imageretrieval/.