GEOMETRICALLY INVARIANT OBJECT-BASED WATERMARKING USING SIFT FEATURE

Viet Quoc PHAM † *Takashi MIYAKI*‡ *Toshihiko YAMASAKI* † *Kiyoharu AIZAWA* †

† Dept. of Information and Communication Engineering ‡ Dept. of Frontier Informatics,
The University of Tokyo
E-mail: {pqvietvn, miyaki, yamasaki, aizawa }@hal.k.u-tokyo.ac.jp

ABSTRACT

In this paper, we have developed a robust object-based watermarking algorithm using the scale-invariant feature transform (SIFT) features in conjunction with a new data embedding method based on Discrete Cosine Transform (DCT). The message is embedded in DCT spaces of randomly generated blocks in the selected object region. To recognize the object region after being distorted, its SIFT features are registered in advance. In the detection scheme, we firstly detect the object region by using feature matching. The transformation parameters are then calculated, and the message can be detected. Experimental results demonstrated that our proposed algorithm is very robust to geometrical distortions such as JPEG compression, scaling, rotation, shearing, aspect ratio change, image filtering, and so on.

Index Terms— Digital watermarking, geometrically invariant, scale-invariant feature transform, object matching

1. INTRODUCTION

Owing to the development of the Internet, digital imaging has experienced tremendous growth over the last decade. We now can easily find and download a large number of images within a few seconds. In order to protect and preserve the owner's right, a number of copyright protection methods have been proposed. Digital watermarking is a technology used for copy control and media identification and tracing. In digital watermarking, they embed a short message (a watermark) in an image or video without affecting the quality but that can be detected using dedicated analysis program.

Due to the advances in image and video editing software, it has been made possible to copy a certain object in an image or a frame and paste it to the others. In addition, such illegally copied object may be geometrically distorted by lossy compression, affine transforms, and so on. The purpose of this paper is embedding and detecting a watermark in such a situation. In this point of view, the scope of this paper is different from conventional object-based watermarking for MPEG-4, in which object layers are pre-defined. In our case, we do not assume such predefined object layers nor do we have to conduct object segmentation for watermarking.

In this paper, we have developed a robust object-based watermarking algorithm using object matching in conjunction with a new data embedding method based on Discrete Cosine Transform (DCT). The general idea of this method is shown in Fig. 1. The watermarked object "akiyo" in Fig. 1 is attacked by being mixed with another object and then geometrically transformed. To detect the hidden information in the object, we first detect the object region by using object matching. And by calculating the affine parameters, we can geometrically recover the object, and can easily read the hidden message. In our method, we employed the SIFT feature [1] for the object matching operation.

The experimental results demonstrated that our method can resist to very strong attacks such as 0.4x scaling, all angle rotation, 30° shearing, JPEG compression (Q=20), StirMark random distortions, or the combination of them.

2. RELATED WORKS

There are several approaches related to geometrically invariant watermarking. We categorize them into two groups.

In the first group, the watermarking systems employ object segmentation. As a representative of this group, Dajun et al. [2] proposed an object-based video authentication system in which a set of angular radial transformation coefficients was selected as the feature to represent the video object and the background. Error correction coding and cryptographic hashing were applied to those selected coefficients to generate the authentication watermark. The watermark embedding and extraction were done by modifying Discrete Fourier Transform (DFT) coefficients. In their detection scheme, the scaling ratio was supposed to be known so that the received video object could be scaled back to its original resolution. Besides, there are some other methods that employ object segmentation, such as [3] [4], etc. One of the most significant disadvantages of such methods is that object segmentation

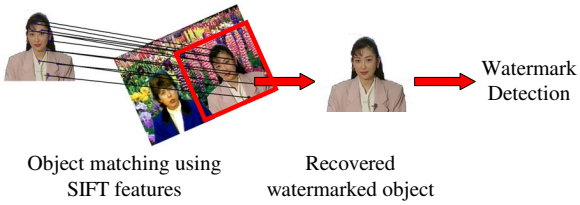


Figure 1: Method demonstration.

Figure 2: Object matching by keypoint pair matching.

generally requires high calculation cost but its reliability is rather low.

In the second group, the watermarking systems embed watermarks in positions located relatively to feature points. Bas et al. [5] proposed an embedding and detection scheme where the mark is bound with a content descriptor defined by salient points. They extracted feature points of the image and perform a Delaunay tessellation on the set of points. The mark was embedded using a classical additive scheme inside each triangle of the tessellation. The detection was done using the correlation properties on the different triangles. With the same idea, Lee et al. [6] extracted circular patches using the scale invariant feature transform (SIFT) [1] descriptor. These circular patches were watermarked additively in the spatial domain. The detection was done using the correlation properties on the different patches like Bas's method [5]. The detection rates of the two methods became lower (see Table 1) for rather strong distortions due to the disappearances of some feature points. Therefore, they repetitively embedded the same watermark pattern in all patches. The correlations between the detected patterns and the reference pattern proved the existence of the watermark. As a result, such methods can be used only in "proof of ownership" applications.

3. SCALE-INVARIANT FEATURE TRANSFORM

Scale-Invariant Feature Transform (SIFT) [1] is an algorithm for extracting distinctive features from images. The algorithm has been used for matching different views of an object or scene and object recognition [7]. The features (called "SIFT features") are invariant to the image scale, rotation, and partially invariant to changing viewpoints, and change in illumination.

The first stage of the computation searches for extrema over all scales and image locations. It is implemented efficiently by using a difference-of-Gaussian (DoG) function to identify potential interest points that are invariant to scale and orientation. Next, keypoints are selected from the candidates based on measures of their stability. Finally, a keypoint descriptor is created by computing the gradient magnitude and orientation at each image sample point in a region around the keypoint location.

As shown in the Fig. 2, two objects are matched by searching the nearest keypoint pairs from the two objects. The nearest keypoint is defined as the keypoint with the minimum Euclidean distance for the invariant descriptor vector.

4. WATERMARKING SCHEME

4.1. Embedding Scheme

As shown in Fig.3, we describe how to embed hidden messages and register information for the detecting scheme. There are three steps in the scheme:

-Step 1: In this step, we select the object region from the original work to embed a hidden message. The object shape can be arbitrary – the bounding rectangle or the segmentation result of the object.

-Step 2: A hidden message is embedded in the selected region. The detailed explanation of the embedding method is described in section 4.3.

-Step 3: In this final step, we extract SIFT features from the object region (Fig. 4) and register them in the database for the object matching in the detection procedure. The position of the selected region is also registered.

4.2. Detecting Scheme

As shown in Fig.5, we describe how to detect the embedded message from the object that was attacked. We divide the detecting scheme in three steps:

-Step 1: We extract the SIFT features from the attacked watermarked image. Then the extracted features will be matched with the registered features (step 3 in section 4.1).

-Step 2: Based on the matching results, we calculate the parameters of the geometrical transformation (six parameters for the affine transformation, or eight parameters for the perspective transformation). In our experiment, we only deal with the affine transformation. The perspective transformation will be solved in the future work.

-Step 3: We recover the attacked object region based on the transformation parameters found in step 2. Then the embedded message can be detected (section 4.3).

4.3. Embedding Method

We embed a sequence of bits in the object by modifying the Discrete Cosine Transform (DCT) coefficients of its region. First, 16x16 blocks are randomly chosen within the object by the secret key *K*. The message is embedded in these blocks. The reason for the random block sampling is to

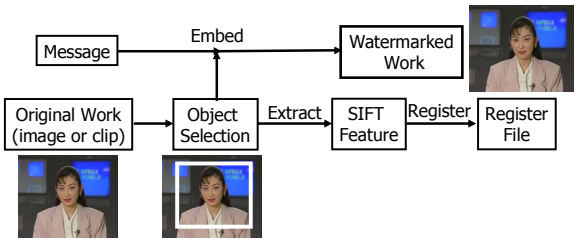


Figure 3: Embedding scheme.

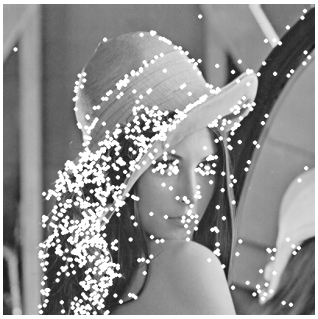


Figure 4: 1082 SIFT feature points are extracted.

prevent attackers from knowing where to attack. After that, two coefficient indices (x_0, y_0), (x_1, y_1) are selected from each 16x16 block. One bit (mark) is then embedded in both (x_i, y_i) $(i = 0, 1)$ by modifying DCT coefficients $f(x_i, y_i)$, $f(y_i, x_i)$

$$(f(x_i, y_i), f(y_i, x_i)) \rightarrow (f'(x_i, y_i), f'(y_i, x_i))$$

When mark = 0:

$$\begin{cases} f'(x_i, y_i) = \dfrac{f(x_i, y_i) + f(y_i, x_i)}{2} - \dfrac{s}{2} \\ f'(y_i, x_i) = \dfrac{f(x_i, y_i) + f(y_i, x_i)}{2} + \dfrac{s}{2} \end{cases}$$

When mark = 1:

$$\begin{cases} f'(x_i, y_i) = \dfrac{f(x_i, y_i) + f(y_i, x_i)}{2} + \dfrac{s}{2} \\ f'(y_i, x_i) = \dfrac{f(x_i, y_i) + f(y_i, x_i)}{2} - \dfrac{s}{2} \end{cases}$$

where s is the watermarking strength.

In the detecting scheme, the embedded mark can be detected by comparing $\Sigma f(x_i, y_i)$ and $\Sigma f(y_i, x_i)$:

If $\Sigma f(x_i, y_i) > \Sigma f(y_i, x_i)$ mark=1, otherwise mark=0.

The two indices (x_0, y_0), (x_1, y_1) are selected from the seven candidates (1,4), (2,3), (1,5), (2,4), (1,6), (2,5), (3,4). In our algorithm, (x_0, y_0), (x_1, y_1) are selected as the two indices that have the smallest value of $|f(x_i, y_i) - f(y_i, x_i)|$.

The strength s is selected by considering the quality of watermarked images. As demonstrated in Fig. 6, the high value of s makes the watermarked image noticeable where as higher s makes the watermark more robust. Based on the empirical study, we set the basis value of s to 0.18. This value will be increased or decreased based on the feature of the local region.

5. EXPERIMENTAL RESULTS

To evaluate our method, we considered a variety of attacks including affine transformations, image filters, background changes, JPEG and MPEG compressions, random distortions (defines in a watermark benchmarking software

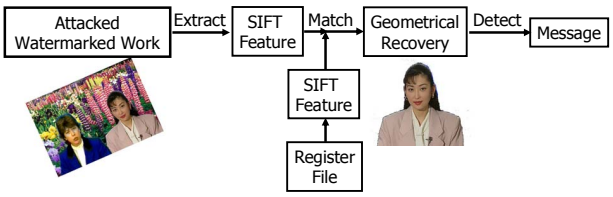


Figure 5: Detecting scheme.

(a) $s = 0.50$ (b) $s = 0.18$

Figure 6: Watermarking strength.

Table 1: Ratio of correctly detected blocks [6].

Attacks	[5]	[6]	Proposed method
JPEG (Q = 40)	70.1%	92.3%	100%
Gaussian 3x3	59.7%	89.3%	100%
Rotation 30°	20.1%	63.7%	100%
Scaling 0.75x	14.2%	62.4%	100%

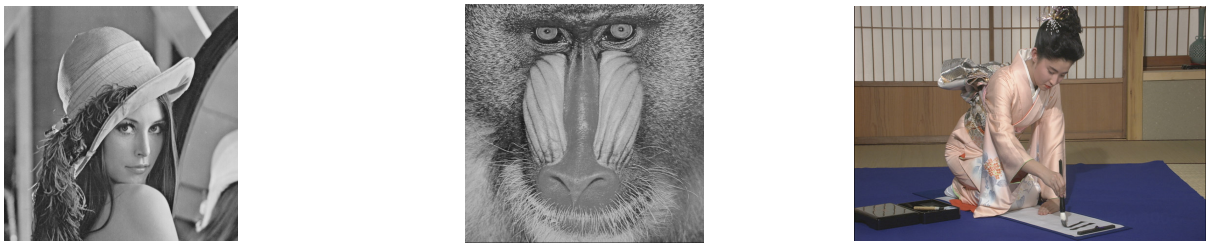
StirMark 4.0 [8]). The samples used in our experiments are "LENA" and "BABOON" for the still images and "Calligraphy Practice" for the video. In case of still image watermarking, we generated 166 blocks within the objects, and embedded 83 bits into them (each bit was embedded redundantly into 2 blocks). In case of video watermarking, 50 bits are embedded redundantly in successive frames.

Figs. 8 (a), (b), and (c) show that for affine transformation attacks, the detection rate can reach up to 97% for 0.4x scaling, nearly 100% for all angle rotations, and 100% for 20° shearing. Our method is also robust to various kinds of image filters and background changes that are common in video editing (Figs. 8 (d) and (e)). For MPEG4 compressions, we repetitively embedded 50 bits in successive frames. The experimental result (Fig 8 (f)) shows that 10 frames are needed for the highest detection rate.

The performance of our algorithm is summarized and compared with [5] [6]. We can see that our approach yields much better results than [5] [6] (Table 1).

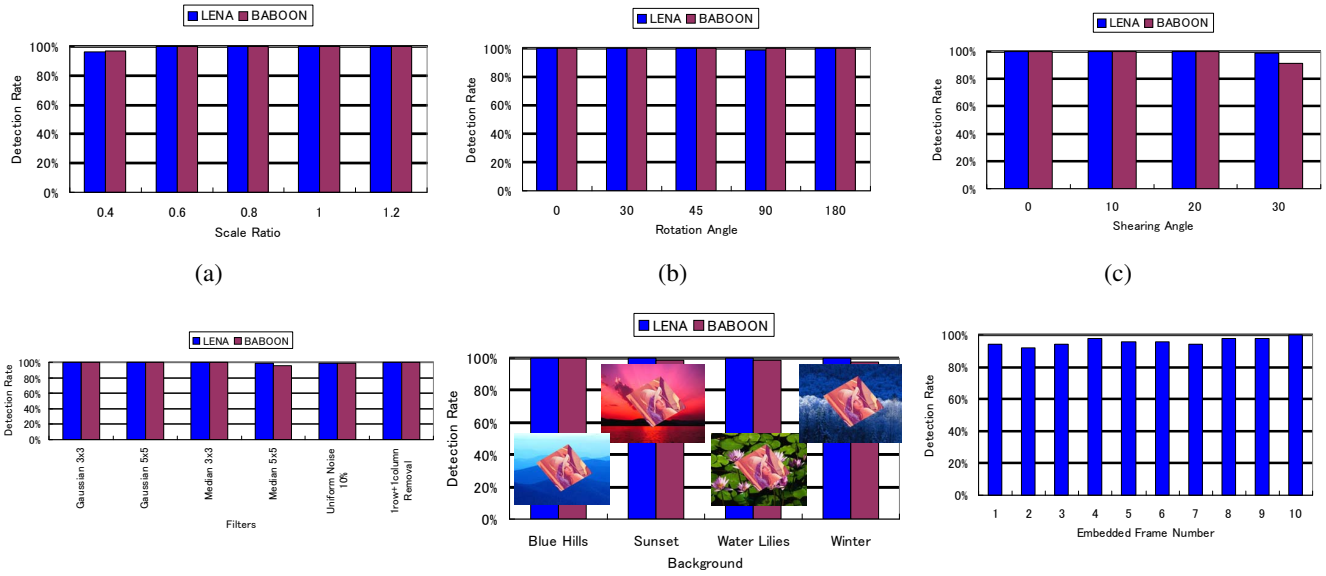
6. CONCLUSIONS

By using the combination of the object matching and the robust watermarking scheme, we have developed a robust geometrically invariant object-based watermarking. The experimental results show that our proposed method can resist to very strong attacks such as 0.4x scaling, all angle



(a) (b) (c)

Figure 7: Images and video used in our experiments: (a) LENA, (b) BABOON, (c) Calligraphy Practice.



(a) (b) (c)

(d) (e) (f)

Figure 8: Resistance against attacks: (a) Scaling attack (fixing rotation angle = 30°, shearing angle = 10°), (b) Rotation attack (fixing scale ratio = 0.8, shearing angle = 10°), (c) Shearing attack (fixing scale radio = 0.8, rotation angle = 15°), (d) Image filters, (e) Background changes, (f) MPEG4.

rotation, 30° shearing, JPEG compression (Q =20) or the combination of all of them. By applying the redundant embedding robustness, we can get much better results for the video watermarking. Its robustness to a wide variety of attacks is suitable for many applications requiring high watermarking reliability and capacity. Our method is an informed watermarking, but the size of the register file is about 40kB which is not a significant problem. Besides, we can improve the speed of our method by replacing SIFT by any faster matching methods such as SURF [9].

7. REFERENCES

[1] D. G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints," IJCV, 60, 2, pp. 91-110, 2004.

[2] H. Dajun, Q. Sun, and Q. Tian, "A secure and robust object-based authentication system," EURASIP J. on Applied Signal Processing, Vol.14, pp.1-14, 2004.

[3] Y. K. Ho and M. Y. Wu, "Robust object-based watermarking scheme via shape self-similarity segmentation," Pattern Recognition Letters, vol.25 no.15, pp.1673-1680, November 2004.

[4] J. S. Lee and W. Y. Kim, "A New Object-Based Image Watermarking Robust to Geometrical Attacks," PCM (2), 2004, pp. 58-64.

[5] P. Bas, J. M. Chassery, and B. Macq, "Geometrically invariant watermarking using feature points," IEEE Trans. on Image Processing, vol. 11 (2002), pp. 1014-1028.

[6] H. Y. Lee, H. Kim, and H. K. Lee, "Robust image watermarking using local invariant features," Optical Engineering 45(3), 037002 (March 2006).

[7] M. Brown, and D. G. Lowe, "Recognising panoramas," ICCV (9), pp. 1218-1225, 2003.

[8] F. A. P. Petitcolas, "Watermarking schemes evaluation," IEEE Trans, Signal Processing, vol. 17, no. 5, pp. 58–64, September 2000.

[9] H. Bay, T. Tuytelaars, L. V. Gool, "SURF: Speeded Up Robust Features," ECCV (9), pp. 404-417, May 2006.