TOWARDS A ROBUST FEATURE-BASED WATERMARKING SCHEME

Jonathan Weinheimer[a], *Xiaojun Qi*[b], *and Ji Qi*[b]

[a]jw010m@mail.rochester.edu
Computer Science Department, University of Rochester, Rochester, NY 14627
[b]xqi@cc.usu.edu and jiqi@cc.usu.edu
Computer Science Department, Utah State University, Logan, UT 84322-4205

ABSTRACT

This paper presents a feature-based watermarking scheme which is robust to geometric and common image processing attacks. The improved Harris corner detector is used to find the robust feature points, which correspond to high corner responses and survive various attacks. The watermark is embedded in the Fourier frequency domain of the disk area centered at each feature point to ensure the resilience to the local attacks and common image processing. The image normalization method is adopted to determine the possible rotation for reducing synchronization errors between the embedded and extracted watermarks. The watermark detection decision is based on the number of matched disks in terms of the number of matched bits between the recovered and embedded watermarks in embedding blocks. Experimental results demonstrate that our scheme is more robust to geometric and common image processing attacks than the peer feature-point-based approaches.

Index Terms – feature-based watermarking, image normalization

1. INTRODUCTION

Watermarking is an important approach to solving tamper proofing and authentication problems associated with the digital media (i.e., audio, video, and images). This paper focuses on image authentication by embedding invisible and robust watermark, which survives any reasonable attack, especially the geometric alterations, to the image.

Geometric attacks are difficult for watermarks to withstand since synchronization errors between extracted and embedded watermarks can be easily magnified. Consequently, invariant-domain-based watermarking [5, 7], template-based watermarking [9], moment-based watermarking [1, 3], and feature-based watermarking [2, 4, 10], have been developed to counterattack geometric distortions. In general, feature-based watermark algorithms are the best approaches to resisting geometric distortions since feature points provide stable references for both watermark embedding and detection. Bas *et al.* [2] use the Harris detector to extract features and Delaunay tessellation to define watermark embedding regions. Kutter *et al.* [4] use the Mexican hat wavelet to extract features and Voronio diagrams to define watermark embedding regions. Tang and Hang [10] also use the Mexican hat wavelet to extract feature points. Watermarks are embedded in the normalized disks centered at the feature points. These approaches embed watermark in the local image content regions which are determined by the detected feature points and therefore are relatively robust against the geometric attacks, local distortions, and cropping. However, the robustness of these methods depends on the capacity of the detector to preserve feature points after geometric transformation, especially on images with more texture and images with less texture and large homogeneous areas.

In this paper, we develop a robust feature-based watermarking scheme to resist geometric distortions and common image processing. It combines the advantages of robust feature extraction, image normalization, and disk correlation to reduce the watermark synchronization errors. The remainder of the paper is as follows. Section 2 describes the proposed robust feature extraction method. Section 3 presents the utility of the image normalization method in restoring the probe image for watermark detection. Section 4 covers the details of our watermark embedding and detection method. Section 5 shows the experimental results. Section 6 draws conclusions.

2. ROBUST FEATURE POINT DETECTION

Feature point detection is an important step in the proposed scheme. In order to detect watermarks without access to the original image, we look for feature points that are perceptually significant and can thus resist various distortions. These feature points can be further used as synchronization markers during the detection process. In our scheme, we use the Harris corner detector for the feature point detection since it ranks the most robust as compared with Archard-Rouquet and SUSAN detectors [2].

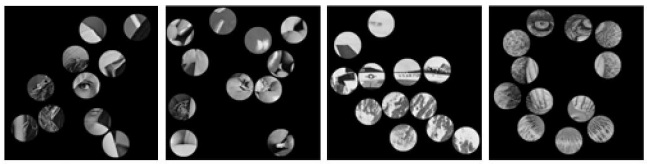
We further improve the Harris corner detector from the following perspectives: 1) Apply a 5×5 averaging filter to smooth the image for alleviating the noise effect. 2)

Calculate the corner response image R by using a unique function [6] as follows:

$$R = \frac{\left(\frac{\partial^2 I}{\partial x^2}\right)\left(\frac{\partial^2 I}{\partial y^2}\right) - \left(\frac{\partial^2 I}{\partial x \partial y}\right)^2}{\left(\frac{\partial^2 I}{\partial x^2}\right) + \left(\frac{\partial^2 I}{\partial y^2}\right)} \quad (1)$$

This function involves the second partial derivatives of the smoothed image I with respect to x and y directions at each position. 3) Find the relative important feature points by keeping all the points whose response values in R are greater than 700, which is an empirically determined threshold for achieving a reasonable number of feature points for most images. 4) Locate the final robust feature points by repeatedly searching for the strongest relative important feature points (i.e., local maxima points) within the non-overlapping disk areas with a radius of 90 pixels.

Fig. 1 demonstrates the final robust feature points by applying our improved Harris corner detector on four images with different textures. These feature points are the centers of the disks shown in Fig. 1. Each disk has a radius of 90 pixels and will be used independently for embedding two watermark sequences.



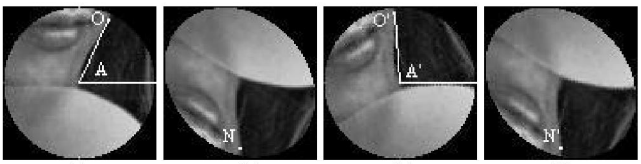
Lena Pepper Plane Baboon

Fig. 1: Robust feature points (centers of the disks)

3. NORMALIZATION FOR SYNCHRONIZATON

The image normalization method [8] has been extensively used in pattern recognition. It transforms the original image to the normalized image by using the central moments, the covariance matrix, and its eigenvalues. This normalized image is invariant to rotation and therefore is a good candidate for embedding watermarks. However, the normalized coordinates are non-integer values which not only require the spatial interpolation before the embedding process but also make the watermark detection difficult. Consequently, we propose to use the image normalization to regain the local synchronization of the watermarking areas between the host and probe images. To this end, we reduce the use of the normalized coordinates to minimize the interpolation errors that occur in this process.

Fig. 2 shows the normalized disks of a sample original disk from Lena image and its 30° rotated disk. The steps of utilizing the image normalization to regain the synchronization between each disk and its geometrically attacked disk are summarized below. Five steps are first applied to the original disk to find the reference angle A. 1) Convert the original disk to the normalized disk as shown in Fig. 2(b) by a normalization transform T_1. 2) Use a secret key K to generate a reference coordinate in the normalized disk. 3) Locate a single normalized coordinate N as shown in Fig. 2(b) that is closest to the given reference coordinate. 4) Apply the inverse normalized transform T_1^{-1} to find the reference spatial coordinate O in the original disk as shown in Fig. 2(a). 5) Calculate the angle A as shown in Fig. 2(a) with respect to the disk center and the horizontal line across the disk center. Apply the above 5 steps to the probe disk as shown in Fig. 2(c) to calculate the angle A' by using the same secret key K. The difference between A and A' (i.e., $\beta = A - A'$) determines the rotated angle. This angle can be used to restore the probe disk so the local synchronization between extracted and embedded watermarks is regained. That is, this angle difference can be used to restore the probe disk to be aligned with the original disk.



(a) (b) (c) (d)

Fig. 2: The illustration of the interpolated normalized disks (a) An original disk from Lena image (b) Its normalized disk (c) Its 30° rotated disk (d) The normalized disk of (c)

4. WATERMARK EMBEDDING AND DETECTION

Fig. 3 illustrates the block diagram of our proposed watermark embedding scheme. The robust feature points are first extracted and a set of non-overlapping disks centered at these feature points is found. Two non-overlapping 32×32 blocks are selected inside each disk. These two blocks are respectively selected above and below the feature point using secret keys. Fourier transformation is then applied to the blocks and secret-key-based 16 points (x_i, y_i)'s and 16 binary watermark bits W_i's are generated for embedding. In detail, the Fourier magnitudes F_A and F_B at points $A = (x_i, y_i)$ and $B = (-y_i, x_i)$ are altered as follows:

$$W_i = 1: \begin{cases} F_A' = F_A + 0.5[\alpha - (F_A - F_B)] \\ F_B' = F_B - 0.5[\alpha - (F_A - F_B)] \end{cases} \quad (2)$$

and

$$W_i = 0: \begin{cases} F_A' = F_A - 0.5[\alpha - (F_A - F_B)] \\ F_B' = F_B + 0.5[\alpha - (F_A - F_B)] \end{cases} \quad (3)$$

where F_A' and F_B' are the new magnitudes at points A and B, and α is the watermark strength which is calculated as:

$$\alpha = 0.5 \times g \times (F_A + F_B) \quad (4)$$

Here g is the gain factor and experimentally set to be 0.8 to ensure a balance between watermark robustness and invisibility. As a result, the watermark strength α is adaptive to the magnitude values at points A and B. Finally, the inverse Fourier transform is performed on the adjusted magnitudes and the restored watermarked blocks replace the original blocks to generate the watermarked image.

Additional information, including the secret keys for generating the block positions, embedding positions, binary

watermark bit sequence, and the reference point for the normalized disk, and the average intensity and angle values for each disk, is saved for detection. The angle is computed by the normalization method explained in Section 3.

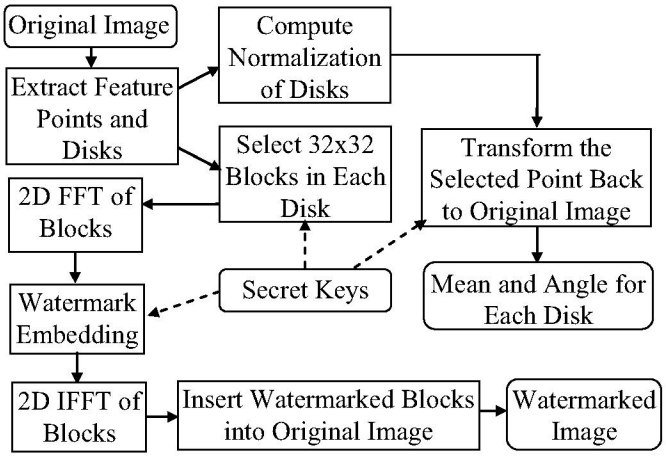


Fig. 3: Watermark embedding process

The watermark detector does not need the original image. Fig. 4 shows the block diagram of our proposed detection scheme. Similar to the embedding procedure, a set of feature-point-centered disks is located. The average intensity values of these disks are compared with the saved average intensity values of the original disks to determine the correlation between each pair. That is, if the average intensity of the probe disk differs that of the saved original disk by 0.5, these two disks are paired. The 5 steps introduced in Section 3 are then applied to each correlated disk pair to find its associated restoration angle β_i. A score function is further defined to determine the possible outlier restoration angles which are resulted from the falsely correlated disks. This score function is:

$$Score(\beta_i) = \sum_{j=1}^{n} f(\beta_i, \beta_j) \tag{5}$$

where $f(\beta_i, \beta_j) = \begin{cases} 1 & \text{if } |\beta_i - \beta_j| < .01 \\ 0 & \text{if } |\beta_i - \beta_j| \geq .01 \end{cases}$. That is, the possible outlier restoration angles should have a low score and the possible correct restoration angles should have a high score. Consequently, the final restoration angle β for the entire image is computed as the mean of the restoration angles with the top highest scores. The probe image is then inversely rotated by β to be aligned with the original image. The feature point detection process is performed again on the restored image to locate the disk centers, which refer to the same content points in both original and probe images. As a result, the local synchronization of each disk is achieved. Afterwards, the same steps prior to the watermark embedding step are followed. The Fourier magnitudes F_A and F_B at points $A = (x_i, y_i)$ and $B = (-y_i, x_i)$ are utilized to extract one embedded watermark bit. That is, if $F_A - F_B < 0$, watermark bit 0 is extracted; otherwise, watermark bit 1 is extracted. This process is repeated for all the 16 positions in each block of each disk to extract the corresponding watermark bit sequence.

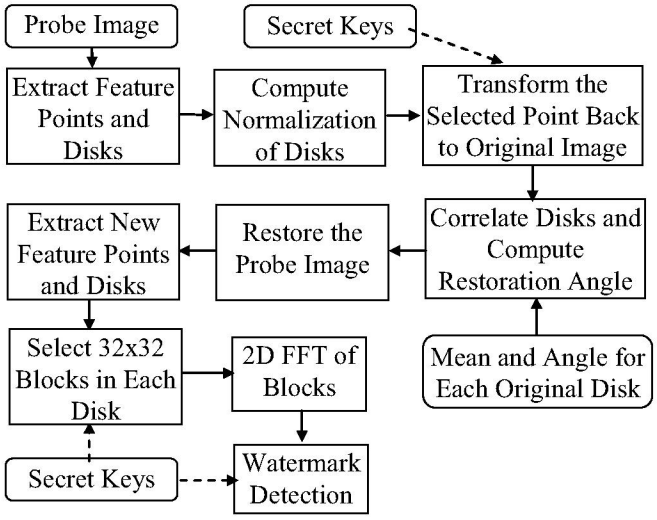


Fig. 4: Watermark detection process

The number of matched bits between the embedded watermark W and the extracted watermark W' determines whether the detected image contains watermark. In general, watermark is present in the detected disk if the total number of matched bits extracted from 2 embedding blocks exceeds T, which is determined based on the false alarm probability:

$$P = \sum_{\substack{k_1=T_1, k_2=T_2 \\ k_1+k_2 \geq T}}^{n} \left(\frac{1}{2}\right)^n \cdot \left(\frac{n!}{k_1!(n-k_1)!}\right) \left(\frac{1}{2}\right)^n \cdot \left(\frac{n!}{k_2!(n-k_2)!}\right) \tag{6}$$

where k_i is the number of matched bits in block i and n is the length of the watermark bit sequence. For instance, if $n = 16$, choosing T as 24 will lead to false alarm probability of $P = 5\times10^{-6}$. In our scheme, we decide the presence of the watermark if at least 2 disks claim to contain watermark.

5. EXPERIMENTAL RESULTS

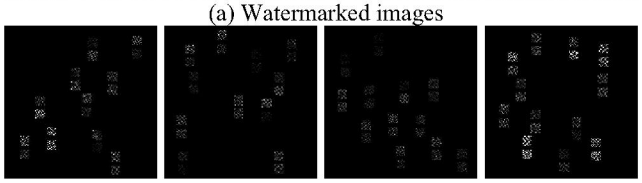
The proposed method has been tested on a variety of images with different textures and different attempting attacks.

Watermark invisibility is evaluated on images of Lena, Peppers, Plane, and Baboon. The PSNR values for these 4 watermarked images are 48.52, 52.34, 49.87, and 45.07, respectively. These values are greater than 35.00db, which is the empirical value for the image without any perceivable degradation. Fig. 5 shows the watermarked images and the difference between the original and watermarked images. It further demonstrates the watermark invisibility.

Simulation results for geometric distortions and common image processing attacks are presented in Tables 1 and 2. The numerical numbers listed in the tables indicate the fraction of correctly detected watermarked disks (i.e., detection rates). To ensure fair comparison, we list the results of the attacks survived by Tang's method [10] in

Table 1. Furthermore, we list 17 attacks that Tang's method cannot handle, namely relatively large rotations and combination attacks in Table 2. Our method outperforms Tang's method under various filters, cropping, and JPEG compression in terms of the detection rates. In addition, our method is robust for a wide variety of rotations and combination attacks whereas Tang's method is only robust to a rotation of less than 3°.

(a) Watermarked images



(b) Difference images (30 times the absolute difference)

Fig. 5: The invisibility of the watermarked images

Table 1: Detection rates under attacks survived by Tang's method

Attacks	Lena		Peppers		Baboon		Plane	
	Ours	Tang	Ours	Tang	Ours	Tang	Ours	Tang
Watermarked Image	10/10	7/8	11/11	4/4	12/12	10/11	12/12	n/a
Median Filter 2x2	6/10	1/8	9/11	1/4	5/12	6/11	9/12	n/a
Median Filter 3x3	6/10	1/8	8/11	1/4	7/12	2/11	8/12	n/a
Sharpening 3x3	6/10	4/8	7/11	4/4	4/12	4/11	5/12	n/a
Gaussian Filter 3x3	7/10	5/8	8/11	1/4	6/12	8/11	8/12	n/a
JPEG 90	10/10	n/a	8/11	n/a	5/12	n/a	6/12	n/a
JPEG 80	4/10	6/8	6/11	3/4	6/12	9/11	6/12	n/a
Cropping 5%	7/10	2/8	5/11	2/4	7/12	2/11	5/12	n/a
Rotate 1°	9/10	n/a	10/11	n/a	7/12	n/a	11/12	n/a
Rotate 1°+Crop 5%	8/10	3/8	5/11	2/4	8/12	3/11	4/12	n/a

Table 2: Detection rates under attacks failed by Tang's method

Attacks	Lena	Peppers	Plane	Baboon
Rotate 2°	7/10	9/11	10/12	9/12
Rotate 4°	6/10	9/11	9/12	6/12
Rotate 5°	8/10	6/11	8/12	6/12
Rotate 10°	5/10	9/11	11/12	4/12
Rotate 15°	3/10	7/11	8/12	5/12
Rotate 20°	3/10	2/11	6/12	6/12
Rotate 30°	5/10	6/11	6/12	4/12
Rotate 60°	1/10	3/11	7/12	2/12
Rotate 90°	7/10	9/11	8/12	7/12
Scaling 0.9	5/10	6/11	6/12	2/12
Rotate 5° + Crop 5%	4/10	1/11	2/12	3/12
Rotate 30° + Crop 5%	4/10	5/11	5/12	2/12
Rotate 60° + Crop 5%	4/10	4/11	2/12	3/12
Rotate 90° + Crop 5%	6/10	5/11	5/12	6/12
Median 2x2+JPEG 90	7/10	6/11	8/12	6/12
Median 3x3+JPEG 90	7/10	7/11	8/12	3/12
Sharp 3x3 + JPEG 90	7/10	9/11	7/12	2/12

Simulation results are also compared with the results yielded from the CBS (Content Based Scheme) [2] by applying Stirmark attacks such as small shearing, rotation 10°, scaling 0.8, and JPEG 50%. The experimental results show that the proposed scheme can successfully detect the watermarks under these attacks and therefore it has comparable performance as the CBS.

6. CONCLUSIONS

In this paper, we propose a watermarking scheme which is robust against most geometric and signal processing attacks. The major contributions are: 1) Use improved Harris corner detector to find robust feature points for re-synchronization. 2) Adopt the image normalization method to restore the probe embedding disk by using one secret-key-based reference point. 3) Combine the advantages of the feature points, disks correlation and image normalization to make the watermark resistant to geometric attacks. 4) Apply embedding in the Fourier domain to make the watermark robust to signal processing attacks.

The proposed method is robust against a wide variety of tests as indicated in the experimental results. In particular, it is more robust against the combination of the geometric distortions than other feature-based watermarking techniques. Our approach can be further improved by developing more reliable feature extraction method under severe geometric distortions and improving the watermark capacity.

7. REFERENCES

[1] M. Alghoniemy and A. H. Tewfik, "Geometric Invariance in Image Watermarking," *IEEE Trans. IP*, Vol. 13, No. 2, pp. 145-153, 2004.

[2] P. Bas, J.M. Chassery, and B. Macq, "Geometrically Invariant Watermarking Using Feature Points," *IEEE Trans. IP*, Vol. 11, No. 9, pp. 1014-1028, 2002.

[3] H.S. Kim and H.K. Lee, "Invariant Image Watermark using Zernike Moments," *IEEE Trans. CSVT*, Vol. 13, No. 8, pp. 766-775, 2003.

[4] M. Kutter, S.K. Bhattacharjee, and T. Ebrehimi, "Toward Second Generation Watermarking Schemes," in *Proc. IEEE ICIP*, Vol. 1, pp. 320-323, 1999.

[5] C.Y. Lin, M. Wu, J.A. Bloom, I.J. Cox, M.L. Miller, and Y.M. Lui, "Rotation, Scale, and Translation Resilient Watermarking for Images," *IEEE Trans IP*, Vol. 10, No. 5, pp. 767-782, 2001.

[6] A. Nobel, "Descriptions of Image Surfaces," PhD thesis, Dept. of Engineering Science, Oxford University, p45, 1989.

[7] J.J.K.O'Ruanaidh and T. Pun, "Rotation, Scale and Translation Invariant Spread Spectrum Digital Image Watermarking," *Signal Processing*, Vol. 66, No. 3, pp. 303-317, 1998.

[8] S.C. Pei and C.N. Lin, "Image Normalization for Pattern Recognition," *Image and Vision Computing*, Vol. 13, No.4, pp. 711-723, 1995.

[9] S. Pereira and T. Pun, "Robust Template Matching for Affine Resistant Image Watermarks," *IEEE Trans. on IP*, vol. 9, no. 6, pp. 1123-1129, 2000.

[10] C.W. Tang and H.M. Hang, "A Feature-Based Robust Digital Image Watermarking Scheme," *IEEE Trans. SP*, Vol. 51, No. 4, pp. 950-959, 2003.