A FEATURE-BASED DIGITAL IMAGE WATERMARKING FOR COPYRIGHT PROTECTION AND CONTENT AUTHENTICATION

Jen-Sheng Tsai, Win-Bin Huang, Chao-Lieh Chen, Yau-Hwang Kuo

Dept. of Computer Science and Information Engineering
National Cheng Kung University, Taiwan, R.O.C.

ABSTRACT

In this paper, a feature-based robust digital image watermarking algorithm is proposed to achieve the goal of image authentication and protection simultaneously. The Hessian-Affine feature detector is at first adopted to extract characteristic regions of an image in the proposed watermark embedding scheme. Then the copyright watermark is embedded into these characteristic regions according to the local orientation of each pixel. Moreover, the remainder regions are applied for image authentication by using block-wise fragile watermarking method. Similarly, the proposed watermark detection scheme follows the above procedure to extract the copyright watermark and the authentic information independently and blindly from watermarked images. Various attacks are also applied to the watermarked images in order to examine the robustness of our algorithm. The experimental results in this paper show that the proposed watermarking algorithm can resist most removal and geometric attacks. Besides, changes or modifications of an image will be reflected in our hidden watermarks.

Index Terms — Digital watermarking, Feature extraction, Copyright protection, Content authentication.

1. INTRODUCTION

Recently, digital watermarking for copyright protection and content authentication has been the most important secure issue in the digital world. In general, digital watermarking is typically classified into two categories: robust and fragile. The robust watermarking is usually applied for copyright protection and ownership verification [1-6] because that aims to keep embedded watermarks recognizable as possible after various attacks. For content authentication, changes or modifications of an image will be reflected in hidden watermarks when the fragile watermarking is applied [7-8]. However, the performance of a robust watermarking usually depends on how many kinds of attacks are resisted and how much information is embedded at a certain image quality. Different attacks are resisted by using different kinds of watermarking methods. In the literature, various attacks are attempted to destroy or invalidate the embedded watermarks, and those are roughly classified into two types, noise-like signal processing and geometric distortions. Noise-like signal processing mainly tries to remove a watermark from cover data, such as compression, de-noising and low-pass filter etc., and many methods are proposed to resist this kind of attacks [17]. Geometrical distortions mostly cause a watermarking detector to fail to detect the existence of the watermark in spite of the watermark still in an image. Similarly, there also have been many kinds of methods for resisting geometric distortions, such as the transform-based scheme [1], the pilot-based scheme [2], and the feature-based scheme [3-6] etc. The transform-based scheme embeds a watermark in affine-invariant domain such as Fourier-Mellin transform. The pilot-based scheme embeds not only an ownership watermark but also a redundant watermark as a pilot signal for re-synchronization. An invariant point extracted from an image by using the feature-based scheme is used to insert an ownership watermark. More discussion on those algorithms can be found in [13]. In this paper, a feature-based watermarking algorithm is proposed, and the purpose of content authentication is also achieved.

Since a robust feature region for resisting various attacks can be obtained by a good feature detector, the feature detector is a key role in the feature-based watermarking. Up to present, many good feature detectors have been proposed such as Hessian-Affine detector [9], Harris-Affine detector [9], MSER [14], IBR [15], and EBR [15] etc. According to the experimental results in [10], the Hessian-Affine and MSER detectors have better performance. However, the MSER detector is difficultly applied to a watermark algorithm because of the irregular output region of the detector. Therefore, the Hessian-Affine detector is adopted in the proposed algorithm to obtain the feature points and the characteristic regions. Moreover, the proposed algorithm then exploits the local orientation of each pixel to embed the copyright watermark into the regions. In order to achieve different kinds of applications simultaneously, the remainder regions are also used for fragile watermarking by applying block-wise methods [8]. The experimental results show that the proposed watermarking algorithm can resist most removal and geometric attacks. Besides, changes or modifications of an image will also be reflected in our hidden watermarks.

The rest of the paper is organized as follows. The Hessian-Affine detector is introduced in section 2. The proposed feature-based robust watermarking algorithm is presented in section 3, and the experimental results are shown in section 4. The concluding remarks are drawn in section 5.

2. THE HESSIAN-AFFINE FEATURE DETECTOR

A scale and affine invariant interest point detector, called Hessian-Affine detector, is proposed by Mikolojczyk and Schmid in [9]. The detector is adopted to obtain the feature points and the characteristic regions of an image in our proposed algorithm. The procedure of the Hessian-Affine detector is as follows:

1) Detect initial points with Hessian detector and select the characteristic scale.
2) Estimate the shape with the second moment matrix.
3) Normalize the ellipse region to circular one.

4) Refine the point location and scale.
5) Go to step 2 if the second moment matrix of new point is not isotropic.

In first step, the Hessian detector is based on the Hessian matrix H.

$$H(\mathbf{x},\sigma_D)=\begin{bmatrix} L_{xx}(\mathbf{x},\sigma_D) & L_{xy}(\mathbf{x},\sigma_D) \\ L_{xy}(\mathbf{x},\sigma_D) & L_{yy}(\mathbf{x},\sigma_D) \end{bmatrix} \quad (1)$$

where $\mathbf{x}$ is the coordinate of a point (x,y), σ_D is the differentiation scale of the Gaussian kernel, and L_{ij} is the second derivative of a point with respect to i and j variables. A point is regarded as a feature point if the second derivative test discriminant of the point is a local maximum. Moreover, a scale selection function, called Laplacian of Gaussian (LoG), in equation 2 is applied in order to deal with scale changes.

$$\left|LoG(\mathbf{x},\sigma_n)\right| = \left|\sigma_n^2(L_{xx}(\mathbf{x},\sigma_n)+L_{yy}(\mathbf{x},\sigma_n))\right| \quad (2)$$

where σ_n indicates the Gaussian scale factor at scale n. The operator responses are computed for a set of scales σ_n. The response attains an extreme when the size of the LoG kernel matches with the size of a blob-like structure.

In step 2, the elliptical region of the feature point is obtained by using the eigenvalues of the second moment matrix, μ, of the point.

$$\mu(\mathbf{x},\sigma_I,\sigma_D)=\sigma_D^2 G(\mathbf{x},\sigma_I)*\begin{bmatrix} L_x^2(\mathbf{x},\sigma_D) & L_xL_y(\mathbf{x},\sigma_D) \\ L_xL_y(\mathbf{x},\sigma_D) & L_y^2(\mathbf{x},\sigma_D) \end{bmatrix} \quad (3)$$

where σ_I is the integration scale.

In step 3, the obtained elliptical region of the feature point then is transformed into circular one according to the square root of the second moment matrix of the point.

In step 4, a new feature point is obtained when the obtained circular region is also applied to the Hessian detector with a new differential scale. Finally, the procedure would be repeated if the eigenvalues of the second moment matrix of the new feature point is not equal to that of the original one.

(a) **(b)**

Fig. 1. The feature points of (a) the Lena image, and (b) the Lena image which has been rotated, scaled, and cropped.

After the Hessian-Affine detector, many feature points and regions are obtained in the processed image, and the example of the Lena image is shown in Fig. 1(a). Furthermore, the feature points of the Lena image which has been rotated 10 degree, scaled with 1.16 times, and cropped with 13.68 % are shown in Fig. 1(b). Most of the obtained feature points in the original image are the same as those in the rotated one. Therefore, the robustness of the feature detector is exploited to resist various attacks in the proposed algorithm.

3. THE PROPOSED WATERMARKING METHOD

3.1. Embedding watermarks

The block diagram of the proposed embedding watermarks scheme is shown in Fig. 2. At first, the feature points of an image are obtained by using the Hessian-Affine detector. Moreover, the robust characteristic regions are chosen by the region selection. Then the following procedure is divided into two parts: the copyright insertion and the authentic insertion. The copyright insertion is mainly embedding the copyright watermark into the selected regions, and the fragile watermark is embedded into the remainder regions by the authentic insertion.

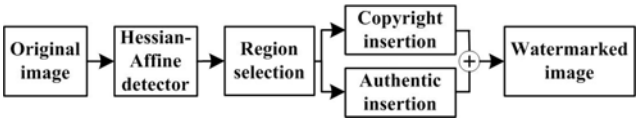

Fig. 2. The block diagram of the watermark embedding scheme.

Many feature points and characteristic regions are obtained after the feature detector. However, some of the feature points are useless and redundant because of the weakness of a region and the overlap between the regions. At first, the region selection removes over large or small regions, because a big or small characteristic region will be vulnerable if the local geometric transform is applied. In our experiments the regions whose characteristic scale is below 2 or above 12 are removed. Moreover, the region with smaller second derivative test discriminant is also removed when there are regions overlapped with each other. An image has been divided into two parts after the region selection, the selected characteristic regions and the remainder regions. The first regions are processed by the copyright insertion, and the others are processed by the authentic insertion. The procedures of the copyright insertion and the authentic insertion are introduced as follows.

The copyright insertion

1) Each selected characteristic region is transformed from an elliptical one to circular one by using the square root of the second moment matrix of a feature point.
2) The direction of the gradient of the pixels within circular region is calculated. The most direction within circular region is defined as the orientation of the circular one. Rotate the orientation of a circular region to constant direction.
3) Noise Visible Function [11] is used to maintain the perceptual quality of an image when watermarks are inserted into the image. The perceptual mask, *NVF*, of a pixel $\mathbf{x}$ is shown as follows:

$$NVF(\mathbf{x})=\frac{1}{1+\theta\cdot s(\mathbf{x})}\ ,\ \theta=\frac{D}{s_{\max}} \quad (4)$$

where $s(\mathbf{x})$ is the local variance of the pixel, and $s_{\max}$ is the maximum local variance of the image, and D is an experimentally determined constant. The D is set for 50 or 100 in our experiment.
4) The sequence of a copyright watermark, $w_c(j)$, is generated by a pseudo-random generator with a secret key. All elements of the sequence are mapped into a bipolar domain $\{-1,+1\}$.
5) The generated copyright watermark is embedded into the circular characteristic region. The embedding function is shown as follows:

$$I_w(\mathbf{x})=I(\mathbf{x})+(1-NVF(\mathbf{x}))\cdot w_c(j)\cdot c_1+NVF(\mathbf{x})\cdot w_c(j)\cdot c_2 \quad (5)$$

where $I(\mathbf{x})$ is the original value of the pixel $\mathbf{x}$, $I_w(\mathbf{x})$ is the value of the pixel embedded with the watermark $w_c(j)$, and c_1 and c_2 are experimentally determined parameters.

6) After embedding the copyright watermark into all selected characteristic regions, each region is translated back to an elliptical one.

The authentic insertion

1) The remainder regions are divided into several 8x8 blocks, and all blocks are indexed. Each block is processed by the following steps 2 ~ 4.
2) Every pixel of a block without the least significant bit is hashed with the width and height of the image and a secret key for authentication, and the block index by the hash function called Rivest's Message Digest version 5 (MD5) [16].
3) A fragile watermark is also generated by a pseudo-random generator with the secret key for authentication which is similar to the step 4 in the copyright insertion.
4) The output sequence of the hash function is exclusive-ORed with the generated fragile watermark. Then the least significant bit-plane of each pixel in a block is replaced by the result.

Finally, we combine the copyright and authentic regions to perform a watermarked image. The watermarked image achieves copyright protection and content authentication simultaneously.

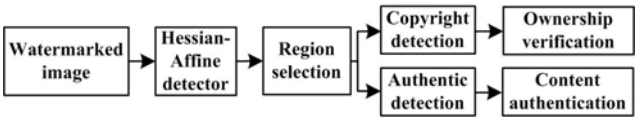

Fig. 3. The block diagram of the watermark detection scheme.

3.2. Detecting watermarks

The proposed detecting scheme shown in Fig. 3 is similar to the above embedding scheme except for copyright detection and authentic detection. Once the characteristic regions are selected, the copyright detection and the authentic detection are used to detect the copyright watermark and the fragile watermark, respectively.

The copyright detection

1) The same as the Step 1 in the copyright insertion.
2) The same as the Step 2 in the copyright insertion.
3) A Wiener filter is used to extract the hidden watermark from a difference image which is calculated between the watermarked image and its Wiener-filtered image.
4) A bit-error is the difference between the extracted watermark and the original one. If the bit-error is lower than a predefined threshold, we address the existence of the copyright watermark.

The authentic detection

1) The same as the Step 1 in the authentic insertion.
2) The same as the Step 2 in the authentic insertion.
3) The output sequence in step 2 is exclusive-ORed with the least significant bit of each pixel in the block. A block is modified if there is a difference between the result and the original fragile watermark.

Therefore, the copyright and the authentication of an image are determined according to the two kinds of the extracted watermarks. A detected error is called false-alarm when there is no watermark embedded but detected having one. For an un-watermarked image an extracted bit is treated as independent random variable with probability 0.5. Here we define $P_{SP-region}$ as the probability of success detection for a characteristic region. Based on the Bernoulli trials, the probability can be

$$P_{SP-region} = \sum_{i=n-T}^{n} \binom{n}{i} \cdot (0.5)^i \cdot (0.5)^{n-i} \tag{6}$$

where T indicates the predefined threshold, and the parameters i is the number of the matching bits and n is the length of watermark bits. We announce that the watermarks are existed in an image if there are at least m regions detected. Therefore, the false-alarm probability, $P_{SP-image}$, of an image can be as:

$$P_{SP-image} = \sum_{j=m}^{N} \binom{N}{j} \cdot (P_{SP-region})^j \cdot (1 - P_{SP-region})^{N-j} \tag{7}$$

where N are total regions found in an image, and j is the numbers of matching regions. In our experiments the average regions found in an image is 20. The above parameters are setting for $T = 30$, $n = 100$, and $N = 20$. Moreover, the false-alarm probabilities of an image for $m = 1, 2, 3$ are 7.8×10^{-4}, 2.9×10^{-7}, and 6.9×10^{-11}, respectively.

4. SIMULATION RESEULTS

Two well-known 512×512 images, Lena and Baboon, are used to evaluate the performance of the proposed algorithm. After embedding invisible copyright and fragile watermarks by using our algorithm, the peak-signal-to-noise-ratio (PSNR) of the Lena and the Baboon images is 45.62 dB and 42.17 dB, respectively. Moreover, two feature-based watermarking methods, Tang's method [4] and Seo's method [5], are compared with our algorithm. All attacks used in the experiment are generated by the benchmark program, StirMark 3.1 [12].

The comparison among the three algorithms is shown in Table 1. The first column indicates the attacks applied to a watermarked image. The 2th ~ 4th columns are the detection ratio of our algorithm, Seo's algorithm, and Tang's algorithm, respectively. Detection ratio refers to the ratio of the number of extracted regions from original images to the number of correctly redetected regions from attacked images. Obviously, the experiment result shows that most of the copyright watermarks after noise-like signal processing and geometric distortions are still detectable by the proposed algorithm. Compared with other methods, the proposed algorithm for resisting most of attacks has better results, the aspect ratio change especially.

(a) (b) (c)

Fig. 4. (a) The watermarked Lena image; (b) The (a) image with some insertions; (c) The result of the tampering detection.

Moreover, the content authentication can also be achieved by using the fragile watermarks which are hidden by the proposed algorithm. For example, the fragile watermarks were embedded into the Lena image shown in Fig. 4(a), and then the watermarked image was covered with some insertions and shown in Fig. 4(b). The result of the tampering detection after the proposed algorithm is shown in Fig. 4(c). It's obviously that the modifications in the processed image are reflected.

5. CONCLUSIONS

A robust feature-based image watermarking algorithm for copyright protection and content authentication is proposed. The Hessian-Affine detector is adopted to obtain the points and regions. Most of the removal and geometric attacks are resisted by our algorithm because of the invariant property of the feature points and characteristic regions. Moreover, the fragile watermarks are also embedded into the non-characteristic regions. Therefore, the robust and fragile watermarks can be detected independently and blindly for different applications. Comparing with other robust watermarking methods, the robustness of the proposed method is superior, and changes or modifications of an image can also be reflected in our hidden watermarks. However, high computing complexity is required for our algorithm since the Hessian-Affine detector is an iterative method. We have to design a faster feature detector whose robustness is the same as the Hessian-Affine detector in the future. Moreover, the accuracy of the location for content authentication is affected since the fragile watermark is not included in the characteristic regions. In order to improve the local precision, the fragile watermarks will be embedded into the characteristic regions in the future.

6. REFERENCES

[1] 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 image," IEEE Trans. Image Processing, Vol. 10, No. 5, pp. 767-782, May 2001.

[2] S. Pereira, and T. Pun, "Robust template matching for affine resistant image watermarks," IEEE Trans. Image Processing, Vol. 9, No.6, pp. 1123-1129, June 2000.

[3] P. Bas, J. M. Chassery, and B. Macq, "Geometrically invariant watermarking using feature points," IEEE Trans. Image Processing, Vol. 11, No. 9, pp. 1014-1028, Sept. 2002.

[4] C. W. Tang and H. M. Hang, "A feature-based robust digital image watermarking scheme," IEEE Trans. Signal Processing, Vol. 51, No. 4, pp. 950-959, April 2003.

[5] J. S. Seo and C. D. Yoo, "Localized image watermarking based on feature points of scale-space representation," Pattern Recognition, Vol. 37, Issue 7, pp. 1365-1375, July 2004.

[6] H. Y. Lee, H. Kim and H. K. Lee, "Robust image watermarking using local invariant features," Journal SPIE, Optical Engineering, Vol. 45, No. 3, March 2006.

[7] E. T. Lin and E. J. Delp, "Review of fragile image watermarks," in Proc. of Multimedia and Security Workshop (ACM Multimedia '99) , pp. 25-29, Oct. 1999.

[8] P. W. Wong, and N. Memon, "Secret and public key image watermarking schemes for image authentication and ownership verification," IEEE Trans. Image Processing, Vol. 10, No. 10, pp. 1593-1601, Oct. 2001.

[9] K. Mikolajczyk and C. Schmid, "Scale and affine invariant interest point detectors," Int. J. Computer Vision, Vol. 60, No. 1, pp. 63-86, Oct. 2004.

[10] K. Mikolajczyk, T. Tuytelaars, C. Schmid, A. Zisserman, J. Matas, F. Schaffalitzky, T. Kadir, and L. V. Gool, "A comparison of affine region detectors," Int. J. Computer Vision, Vol.65, No 1-2, pp. 43-72, Nov. 2005.

[11] S. Voloshynovskiy, A. Herrigel, N. Baumgartner, and T. Pun, "A stochastic approach to content adaptive digital image watermarking," in Proc. 3rd Int. Workshop Information Hiding, pp. 211-236, Sept. 1999.

[12] Fabien A. P. Petitcolas, "Watermarking schemes evaluation," IEEE Signal Processing, Vol. 17, No. 5, pp. 58–64, Sept. 2000.

[13] V. Licks, R. Jordan, "Geometric Attacks on Image Watermarking Systems," IEEE Multi-Media, Vol. 12, No. 3, pp. 68-78, Jul-Sept, 2005.

[14] J. Matas, O. Chum, M. Urban, and T. Pajdla, "Robust wide baseline stereo from maximally stable extremal regions," in Proc. of the British Machine Vision Conference, pp. 384-393, 2002.

[15] T. Tuytelaars and L. Van Gool, "Matching widely separated views based on affine invariant regions," Int. J. Computer Vision, Vol. 59, No. 1, pp. 61-85, August 2004.

[16] R. L. Rivest, "The MD5 message digest algorithm," Tech. Rep., 1992.

[17] I. J. Cox, M. L. Miller, and J. A. Bloom, Digital Watermarking San Francisco, CA: Morgan Kaufman, 2001.

Table 1. The comparison among the three algorithms, the proposed method (Our), Tang and Hang's method (Tang's) [4], and Seo and Yoo's method (Seo's) [5]. The symbol, --, means no data.

	Lena			Baboon		
Attack	**Our**	**Seo's**	**Tang's**	**Our**	**Seo's**	**Tang's**
None	19/19	--	7/8	23/25	--	10/11
Rotation 1	5/19	--	3/8	6/25	--	3/11
Rotation 5	5/19	--	0/8	4/25	--	0/11
Rotation 45	2/19	2/7	--	2/25	1/7	--
Rotation scale 1	3/19	--	0/8	9/25	--	4/11
Rotation scale 45	1/19	--	--	0/25	--	--
Cropping 10%	8/19	--	2/8	8/25	--	2/11
Cropping 25%	5/19	4/7	--	3/25	1/7	--
Cropping 50%	2/19	--	--	2/25	--	--
Linear transform (1.007, 0.010, 0.010, 1.012)	5/19	6/7	5/8	12/25	3/7	4/11
Linear transform (1.010, 0.013, 0.009, 1.011)	9/19	7/7	4/8	14/25	1/7	4/11
Linear transform (1.013, 0.008, 0.011, 1.008)	8/19	7/7	4/8	10/25	0/7	5/11
Aspect Ratio change(0.8, 1.0)	1/19	--	0/8	2/25	--	0/11
Aspect Ratio change(0.9, 1.0)	1/19	--	0/8	4/25	--	0/11
Aspect Ratio change(1.0, 1.1)	2/19	--	0/8	8/25	--	0/11
Aspect Ratio change(1.0, 1.2)	2/19	--	0/8	2/25	--	0/11
Scale 50%	1/19	2/7	--	0/25	0/7	--
Scale 90%	2/19	4/7	--	2/25	2/7	--
Scale 110%	4/19	--	--	4/25	--	--
Shearing x y 5%	2/19	1/7	1/8	5/25	0/7	2/11
Median 2x2	1/19	--	1/8	3/25	--	6/11
Median 3x3	1/19	--	1/8	3/25	--	2/11
Median 4x4	1/19	5/7	--	1/25	1/7	--
Gaussian 3x3	3/19	3/7	5/8	5/25	0/7	8/11
Sharpening 3x3	2/19	1/7	4/8	4/25	0/7	4/11
JPEG 20	0/19	--	--	1/25	--	--
JPEG 40	1/19	1/7	3/8	2/25	1/7	5/11
JPEG 60	2/19	3/7	6/8	4/25	1/7	7/11
JPEG 80	3/19	--	6/8	5/25	--	9/11
Random bending	4/19	4/7	--	0/25	0/7	--
1row and 5 column remove	6/19	--	3/8	9/25	--	6/11
5 row and 17 column remove	3/19	5/7	0/8	9/25	1/7	3/11