Digital Image Watermarking Against Desynchronization Attacks

M Selin

Department of Computer Science, KMEA Engineering College, Edathala

Aluva, Kerala

ABSTRACT

Fast and massive dissemination of image data across the Internet imposes great challenges of protecting images against illegal access and unauthorized reproduction. Image watermarking provides a powerful solution for intellectual protection. This paper presents a new feature-based image watermarking scheme which is robust to desynchronization attacks. The Harris -Laplace detector is used to extract the robust feature points, which can survive various signal processing and affine transformation. A local characteristic region (LCR), is constructed based on the scale-space representation of an image is considered for watermarking. At each LCR, the digital watermark is embedded, by modulating the magnitudes of Discrete Cosine Transform coefficients. The performance of watermark detection is computed based on the correlation coefficient. The correlation coefficient is computed between the embedded watermark bits and the detected bits. The results show that the proposed scheme is invisible and robust against various attacks which include common signals processing and desynchronization attacks.

Key words- Image watermarking, desynchronization attacks, feature points, local characteristic region (LCR).

1. INTRODUCTION

Fast and massive dissemination of image data across the Internet imposes great challenges of protecting images against illegal access and unauthorized reproduction. As an effective and efficient solution, image watermarking superimposes a copyright message into a host image before dissemination and then unauthorized reproduction can be recognized by extracting the copyright information [4]. On the other hand, attacks against watermarking systems have become more sophisticated. In general, these attacks can be categorized into common signal processing, such as lossy compression, low pass filtering, and noise addition, etc. and desynchronization attacks, such as rotation, scaling, translation (RST), random bend attack (RBA), and cropping, etc.

Most of the watermarking schemes are robust to common signal processing attacks, but shows severe problems to desynchronization attacks. The schemes that can overcome desynchronisation attacks can be roughly divided into invariant transform, template insertion, and feature-based algorithms.

Invariant Transform: Here, [5]-[7] the watermark is embedded in an affine-invariant domain by using Fourier–

Mellin transformation, generalized radon transformation, and Zernike moment, respectively. Despite that they are robust against global affine transformations, those techniques involving invariant domain suffer from implementation issues and are vulnerable to cropping and RBA.

Template Insertion: Another solution to cope with desynchronization attacks is to identify the transformation by retrieving artificially embedded references. In [8] and [9], the template is embedded in the discrete Fourier transform (DFT)

domain as local peaks in predefined positions. The embedded local peaks are searched during the watermark detection process in order to yield information about the affine transformations that the image has undergone. However, this kind of approach can be tampered by the malicious attack since anyone can access the peaks in the DFT and easily eliminate them.

Feature Based: The last category is based on media features and this paper belongs to this category. Its basic idea is that by binding the watermark with the geometrically invariant image features, the watermark detection can be conducted without synchronization error. In [10], the Mexican hat wavelet is used to extract features and Voronio diagrams to define local characteristic regions (LCRs) for watermark embedding and detection. The feature based approaches are better than others in terms of robustness.

In this paper, a robust feature-based watermarking scheme is developed. First, the Harris –Laplace detector, which is robust to the image modification, is utilized to extract the feature points. Then, the LCRs are constructed based on the scale space theory [1]. Then, several copies of the digital watermark are embedded into the non overlapped LCRs by modulating the magnitudes of DCT coefficients. By binding watermark with LCR, resilience against desynchronization attacks can be readily obtained. The results show that the proposed scheme is invisible and invariant to common signal processing and desynchronisation attacks.

The paper is organized as follows. A new LCR construction method is described in section 2 .Section 3 covers the watermark embedding procedure. Section 4 contains the details of the watermark detection procedure. The results and performance of the proposed scheme is shown in section 5. And finally, section 6 includes conclusion.

2. LOCAL CHARACTERISTIC REGION CONSTRUCTION BASED ON SCALE SPACE FEATURE POINTS

To develop a feature based synchronization method, image characteristics appropriate for watermarking should be carefully selected. Feature points can act as a mark for location, resynchronization between the watermark embedding and detection [2], and must be robust against various types of common signal processing and geometric distortions. The Mexican Hat wavelet[13] and Harris detector[14] are widely used for watermarking purposes[10]-[12]. The Mexican Hat wavelet is stable under noise-like processing, but it is sensitive to some affine transformations[15]. The Harris detector is stable under majority attacks, such as rotation, translation, and noise addition, but it hardly survives under scaling distortion [16]. To solve these drawbacks, the Harris-Laplace detector is proposed by Mikolaiczyk and Schmid [17] and proved to be invariant to image rotation, scaling, translation, partial illumination changes, and projective transform. Therefore, this method is adopted to extract feature points. However, a feature point provides only the position

information. The watermark embedding and detection should be performed over the self-adaptive image paths which are called LCRs. To cope with this problem, a new LCRs construction method is proposed as an effort to find correspondences between images in which there are large changes in scale. The details of the Harris-Laplace detector and LCRs construction method are described in the next subsections.

2.1 Harris-Laplace detector

The Harris-Laplace detector relies heavily on both the Harris measure and a Gaussian scale-space representation.

A. Harris corner measure

The Harris corner detector algorithm [14] relies on a central principle: at a corner, the image intensity will change largely in multiple directions. This can alternatively be formulated by examining the changes of intensity due to shifts in a local window. Around a corner point, the image intensity will change greatly when the window is shifted in an arbitrary direction. Following this intuition and through a clever decomposition, the Harris detector uses the second moment matrix as the basis of its corner decisions. The Harris corner detector is based on the local auto-correlation function of a signal; where the local auto-correlation function measures the local changes of the signal with patches shifted by a

small amount in different directions. The matrix A , has also been called the autocorrelation matrix and has values closely related to the derivatives of image intensity.

$$A(x) = \sum_{x,y} w(x,y) \begin{bmatrix} I_x^2(X) & I_x I_y(X) \\ I_x I_y(X) & I_y^2(X) \end{bmatrix}$$
$$I_x = I_y$$

where and are the respective derivatives (of pixel intensity) in the x and y direction. The off-diagonal entries are the product l_x l_y

of and , while the diagonal entries are squares of the respective derivatives. The weighting function w(x,y) can be uniform, but is more typically an isotropic, circular Gaussian,

$$w(x, y) = g(x, y, \sigma) = \frac{1}{2\pi\sigma}e^{\left(-\frac{x^2+y^2}{2\sigma}\right)}$$

that acts to average in a local region while weighting those values near the center more heavily.

This A matrix describes the shape of the autocorrelation measure as due to shifts in window location. Thus, if we let λ_1

and λ_2 be the eigenvalues of A, then these values will provide a quantitative description of the how the autocorrelation measure changes in space: its principal curvatures. The A matrix centered on corner points will have two large, positive eigenvalues, and would thus have a large Harris measure.

$$R = \lambda_1 \lambda_2 - \alpha (\lambda_1 + \lambda_2)^2$$

Thus, corner points are identified as local maxima of the Harris measure that are above a specified threshold.

$$\{x_{c}\} = \{x_{c} | R(x_{c}), > R(x_{i}), \forall x_{i} \in W(x_{c})\},\$$

$$R(x_c) > t_{threshold}$$

where $\{x_c\}$ are the set of all corner points, R(x) is the Harris measure calculated at x, $W(x_c)$ is an 8-neighbor set centered around x_c and $t_{threshold}$ is a specified threshold.

B. Gaussian Scale-space

The scale-space representation is a set of images represented at different levels of resolutions. A Gaussian scale-space representation of an image is the set of images that result from convolving a Gaussian kernel of various sizes with the original image. In general, the representation can be formulated as:

$$L(x, y, \delta_D) = G(x, y, \delta_D) * f$$

where G is the associated uniform Gaussian kernel with standard deviation δ_D and mean zero, f is an image and * denotes linear convolution. The convolution with a Gaussian kernel smoothes the image using a window corresponds to the size of the kernel. A larger scale corresponds to a smoother resultant image. For a given image f(x, y), its linear (Gaussian) scale-space representation is a family of derived signals, L(x, y; t) defined by the convolution of f(x, y) with the Gaussian kernel,

$$g(x, y; t) = \frac{1}{2\pi t} e^{-(x^2 + y^2)/2t}$$

such that,

$$L(.,.;t) = g(.,.;t) * f(.,.),$$

where the semicolon in the argument of L implies that the convolution is performed only over the variables x and y, while the scale parameter t after the semicolon just indicates which scale level is being defined.

C. Scale Adapted Harris Detector

The Harris-Laplace detector combines the traditional 2D Harris corner detector with the idea of a Gaussian scale-space representation in order to create a scale-invariant detector. Harris-corner points are good starting points because they have been shown to have good rotational and illumination invariance in addition to identifying the interesting points of the image. However, the points are not scale invariant and thus the second-moment matrix must be modified to reflect a scale-invariant property. The scale adapted second moment matrix is given

$$M(x, y, \delta_I, \delta_D) = \delta^2{}_D.G(\delta_I) * \begin{bmatrix} L^2{}_x(x, y, \delta_D) & L_x L_y(x, y, \delta_D) \\ L_x L_y(x, y, \delta_D) & L^2{}_y(x, y, \delta_D) \end{bmatrix}$$
(1)

where δ_I and δ_D are the integration scale and differentiation scale respectively, and L_{α} is the derivative computed in α direction.

Given the δ_D , the uniform Gaussian scale-space representation L is defined by

$$L(x, y, \delta_D) = G(x, y, \delta_D) * f$$
⁽²⁾

where G is the associated uniform Gaussian kernel with standard

deviation δ_D and mean zero f is an image and * denotes linear convolution. Given δ_I and δ_D , the scale adapted Harris corner

strength (SHCS) can be computed by using the determinant and trace of the second moment matrix.

$$R(x, y, \delta_{I}, \delta_{D}) = Det(M(x, y, \delta_{I}, \delta_{D})) - k Tr^{2}(M(x, y, \delta_{I}, \delta_{D}))$$
(3)

where $Det(\mathbf{I})$ and $Tr(\mathbf{I})$ denote computing the determinant and

the trace of matrix, respectively..At each level of the scale space, the feature points are extracted as the local maxima in the image plane as follows:

$$R(x, y, \delta_{l}, \delta_{D}) > R(\hat{x}, \hat{y}, \delta_{l}, \delta_{D}) \quad \forall (\hat{x}, \hat{y}) \in A$$
(4)

$$R(x, y, \delta_I, \delta_D) \ge t_u$$

A and t_{u} are the neighborhood of point (x, y) and the threshold

where

respectively.

Building upon this scale-adapted second-moment matrix, the Harris-Laplace detector is a twofold process: applying the Harris corner detector at multiple scales and automatically choosing the *characteristic scale*.

D. Automatic Scale and Scale –invariant Feature points

The idea of automatic scale selection is to select the characteristic scale of the local structure, for which a given function attains an extremum over scales. The characteristic scale reflects the maximum similarity between the feature extraction operator and the local image structure[17]. This scale estimate will obey perfect scale invariance under rescaling of an image pattern. Here,

Laplacian-of-Gaussians (LOG) is used to find the characteristic scale. The LOG is defined by

$$LOG(x, y, \delta_{I}) = \delta_{I}^{2} \left| \frac{\partial^{2} G(x, y, \delta_{I})}{\partial x^{2}} + \frac{\partial^{2} G(x, y, \delta_{I})}{\partial y^{2}} \right|$$
(5)

Given a point in the image and a set of scale, the characteristic scale is the at which the LOG attains a local maximum over scale[3].

The extraction of feature points using the Harris-Laplace detector consists of the following two steps.

Step I. We first build a scale-space representation with the Harris function for preselected scales $\delta_n = 1.4\varepsilon^n$, where ε is the scale factor between successive levels (set to 1.4 [18]). At each level of the representation, the SHCS is computed with the integration scale $\delta_I = \delta_n$ and the local scale $\delta_I = s\delta_n$, where *s* is a constant (set to 0.7[17]). We then extract the candidate points which are the maxima in the 8-neighborhood and their SHCSs are higher than the threshold $t_{\mu} = 1000$.

Step II. For each candidate point, we apply an iterative method to determine the location and the scale of the feature points. The extrema over scale of the LOG is used to select the scale of feature points. Given an initial point p with scale δ_1 , the iteration steps are as follows.

Step 1. Find the local extremum over scale of LOG for the point \mathcal{P}_k ; otherwise ,reject the point. The investigated range of scales is limited to $\delta_l^{(k+1)} = t$. $\delta_l^{(k)}$ with $t \in [0.7, \dots, 1.4]$.

Step 2. Detect the spatial point p_{k+1} of a maximum of the SHCS closest to p_k for the selected $\delta_l^{(k+1)}$

Step 3. Go to Step 1, if $\delta_I^{(k+1)} = \delta_I^{(k)}$ or $p_{k+1} = p_k$

E. Local Characteristic Region

LCRs are the subsets of the host image and are used for watermark embedding and detection. The watermark embedding and detection are done on the specialized regions of host images called as the local characteristic region or LCR. Therefore, the problem of geometric synchronization must be considered by the LCR construction method. Here, a new construction method based on characteristic scale is proposed.

The LCRs can be of any shape such as triangle, rectangle and circle. However, it is necessary to assure that the LCRs are invariant to rotation, so the circle area is selected as LCR. Moreover, the size of the LCR should vary with the image scale in order to resist the scaling distortion, so the characteristic scale is helpful to determine the size of the LCRs. The radius R of the LCRs is defined as

$$R = \tau. [\delta] \tag{6}$$

where R is the radius of the LCRs, δ is the characteristic scale, and τ is a positive integer which is used to adjust the size of the LCR. We select LCRs that does not interfere with each other.

F. WATERMARK EMBEDDING SCHEME

All LCRs can be considered as independent communication channels. To improve the robustness of transmitted information (watermark), all channels carry the same copy of the chosen watermark. The information passing through each channel may be disturbed by different types of the intentional and unintentional attacks. During the detection process, we claim the existence of watermark if at least two copies of the embedded watermark are correctly detected.

Different steps in watermark embedding procedure are:

Step 1) Generating the watermark.

A random sequence is $W = \{w_i, i = 1 \dots L\}$ generated by the secret key K_1 , where L is the size of the sequence. The sequence values belong to the set $\{-1, 1\}$ and the mean of sequence is zero.

Step 2) Extracting the LCRs. The Harris–Laplace detector is applied to the host image, and a set of feature points, denoted as P, is obtained .Then, a set of LCRs, denoted as O, is constructed in accordance with the location and characteristic scales of the feature points.

Step 3) Embedding the watermark in the LCRs.

The watermark is embedded in the DCT domain of the LCRs. But it is very hard to perform the DCT on the circle areas, so a zeropadding operation is considered to solve this problem. The circle areas are mapped to the blocks of size 2R x 2R by using zeropadding method, where is R the radius of the circle areas. After zero-padding, a 2-D DCT is then applied to all blocks. The watermarks generated as random sequence are embedded into all blocks. Here, in each block first five DCT coefficients are selected in a zigzag scan, similar to those in JPEG compression .Thus five random numbers generated as watermark are multiplied by a scaling factor, and each of this value is then added to the values of each five coefficients respectively. Thus, the watermarks are embedded into each block. Then 2-D IDCT is taken to convert them back to the spatial domain. After watermark embedding, the zero padded should be removed from these blocks.



Fig.1 ZERO PADDING AND ZERO REMOVING OPERATION

After the aforementioned procedure, the watermarked blocks are mapped to circle area (LCRs) by using zero-removing to replace the original LCRs. As a result, we gain the watermarked image.

G. WATERMARK DETECTION SCHEME

The detection process uses the results of feature point extraction and consequently performs a self-synchronization of the watermark. If the watermarked image undergoes an affine transformation, the set of salient points mainly follows the transformation and several LCRs are consequently conserved. We now detail the following different steps of the detection scheme.

- Step 1) Generating the watermark. The original watermark $W=\{w_i, i = 1 \dots L\}$ is generated depending on the secret key K_1 .
- Step 2) Extracting the LCRs. A set of feature points and a set of LCRs are generated from the watermarked image by using the way that is similar to the embedding scheme.
- Step 3) Extracting the LCRs of original image A set of feature points and a set of LCRs are generated from the original image .
- Step 4) Detecting the watermark in the LCRs.
 Similar to the embedding scheme, the LCRs of the original image and the watermarked should be mapped to the blocks of size 2R X 2R by using a zero-padding operation and a 2-D DCT is then applied to these blocks.

For each blocks of the watermarked image and the original image, the first five DCT coefficients are selected in a zigzag scan, similar to those in JPEG compression. Then the difference of these first five coefficients of the watermarked image and the original image is taken, called the detected watermarks. Correlation coefficient is computed between the embedded watermarks and the detected watermarks. If the correlation coefficient is greater than 0.3, then the watermark is detected. The final detection is claimed "success" when at least two watermarks are detected; otherwise, it "fails."

TABLE 1 DETECTION RATES UNDER COMMON SIGNAL PROCESSING ATTACKS

Attacks	Lena	Barbara	Mandrill	Pirates	Cameraman
Translation x-10 y-10	0.8	0.85	0.97	0.94	1.00
Rotation 5 + Scaling 0.9	0.07	0.17	0.23	0.36	0.41
Translation x-10 y-10 + Rotation 5 + Scaling 0.9	0.23	0.17	0.54	0.38	0.53
Cropping 50%	0.5	0.53	0.76	0.58	0.76

H. RESULT AND ANALYSIS

The proposed watermarking scheme is tested on popular images of Lena (512 X512), Mandrill, Pirates, Barbara and cameraman. The watermarked images are subjected to common signal processing attacks such as median filtering, noise addition, jpeg compression and desynchronization attacks such as rotation, scaling, translation, cropping, and random bend attacks. The watermark detection results show that, the proposed scheme is robust to signal processing and desynchronization attacks. Tables I, II and III summarize the detection results against these attacks. These tables show the detection rates, which is defined as the ratio between the number of correctly detected watermarked LCRs and the number of original embedded watermarked LCRs.

Table 1 shows the ratio of correctly detected watermark to the original, called the detection rates of images such as Lena, Barbara, Mandrill, Pirates and Cameraman which undergoes common signal processing attacks such as low pass filtering, noise addition and Jpeg compression.

Table II shows the ratio of correctly detected watermark to the original called the detection rates of images such as Lena, Barbara, Mandrill, Pirates and Cameraman which undergoes desynchronization attacks such as rotation and scaling.

TABLE II DETECTION RATES UNDER ROTATION AND SCALING

Attacks		Le na	Barbar a	Man dril l	Pirate s	Camera man
Median Filter		0.4	0.67	0.85	0.67	0.41
Gaussian Noise		0.5	0.83	0.89	0.65	0.53
Jpeg Compressio n	70	0.2 7	0.33	0.84	0.35	0.71
	50	0.2 7	0.31	0.72	0.31	0.65
	30	0.2 7	0.25	0.67	0.31	0.59

Table III shows the ratio of correctly detected watermark to the original called the detection rates of images such as Lena, Barbara, Mandrill, Pirates and Cameraman which undergoes desynchronization attacks such as translation, random bend attacks and cropping.

TABLE III DETECTION RATES UNDER TRANSLATION, CROPPING AND RB ATTACKS

Attack	s	Lena	Barba ra	Man dril l	Pirates	Camera man
Rotation	5°	0.2	0.33	0.52	0.63	0.65
	15°	0.2	0.25	0.54	0.52	0.53
	30 °	0.07	0.33	0.51	0.62	0.24
Scaling	0.6	0.07	0.06	0.05	0.09	0.05
	0.9	0.4	0.44	0.71	0.44	0.41
	1.4	0.97	0.85	1.00	1.00	1.00

3. CONCLUSION

Most of the watermarking schemes are robust to common signal processing, but shows severe problems to desynchronization attacks. Based on scale-space theory, a robust image watermarking scheme is designed to survive both common signal-processing and desynchronization attacks. The feature points extracted using the Harris-Laplace detector are reliable under various attacks. It is helpful for resynchronization between watermark embedding and detection. Based on scale-space theory, a size adapted LCRs' construction method is developed, which is effective to resist the scaling attack. The watermark is embedded in the non overlapped LCR's. Watermark detection is done on images that have undergone attacks like noise addition, low pass filtering, Jpeg compression, rotation, scaling, translation, cropping and random bend attacks. Watermark detection is based on the value of correlation coefficient between the embedded bits and the detected bits. If the value of correlation coefficient is greater than 0.3, then watermark detection is claimed to be success. The proposed method is tested on images such as Lena, Barbara, Pirates, Mandrill and Cameraman. The results show that the scheme can overcome common signal processing and desynchronization attacks.

The performance of the proposed scheme could be further improved if the LCR was mapped to the geometrically invariant space. Thus, one of the future researches may be applying some geometrically invariant transformations, such as image normalization. In this way, further robustness to attacks may be achieved.

4. REFERENCES

- [1] Xiangyang Xiangyang Wang, Jun Wu, and Panpan Niu," A New Digital Image Watermarking Algorithm Resilient to Desynchronization Attacks", *IEEE Trans. On Information Forensics and Security*, vol. 2, no. 4, DEC 2007.
- [2] Xiang-Yang Wang, Jun Wu," A Feature-based Robust Digital Image Watermarking Against Desynchronization Attacks," International Journal of Automation and Computing 04(4), October 2007
- [3] Xiang-yangWang, Li-min Hou and Hong-ying Yang,"A feature-based image watermarking scheme robust to local

geometrical Distortions, " J. Opt. A: Pure Appl. Opt. 11 (2009).

- [4] Shiyan Hu," Geometric-Invariant Image Watermarking by Key-Dependent Triangulation," Informatica 32 (2008) ,pp. 169-181..
- [5] J. Ruanaidh and T. Pun, "Rotation scale and translation invariant digital image watermarking," *Signal Process.*, vol. 66, no. 3, pp. 303–317,1998.
- [6] P. Dong, J. G. Brankov, N. P. Galatsanos, Y. Yang, and F. Davoine, "Digital watermarking robust to geometric distortions," *IEEE Trans .Image Process.*, vol. 14, no. 12, pp. 2140–2150, Dec. 2005.
- [7] H. S. Kim and H. K. Lee, "Invariant image watermarking using Zernike moments," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 8, pp.766–775, Aug. 2003.
- [8] S. Pereira and T. Pun, "Robust template matching for affine resistant image watermarks," *IEEE Trans. Image Process.*, vol. 9, no. 6, pp.1123–1129, Jun. 2000.
- [9] X. Kang, J. Huang, Y. Q. Shi, and Y. Lin, "A DWT-DFT composite watermarking scheme robust to both affine transform and JPEG compression," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 8, pp.776–786, Aug. 2003.
- [10] M. Kutter, S. K. Bhattacharjee, and T. Ebrahimi, "Toward second generation watermarking schemes," in *Proc. IEEE Int. Conf. Image Process.*, Kobe, Japan, 1999, vol. 1, pp. 320– 323.

- [11] P. Bas, J. M. Chassery, and B. Macq, "Geometrically invariant watermarking using feature points," *IEEE Trans. Signal Process.*, vol. 11,no. 9, pp. 1014–1028, Nov. 2002.
- [12] C. W. Tang and H. M. Hang, "A feature-based robust digital image watermarking scheme," *IEEE Trans. Signal Process.*, vol. 51, no. 4, pp. 950–958, Apr. 2003.
- [13] B. S. Manjunath, C. Shekhar, and R. Chellappa, "A new approach to image feature detection with applications," *Pattern Recognition.*, vol. 4, pp. 627–640, 1996.
- [14] 14. C. Harris and M. Stephen, "A combined corner and edge detector," in *Proc. 4th Alvey Vision Conf.*, 1988, pp. 147– 151.
- [15] H.-Y. Lee *et al.*, "Evaluation of feature extraction techniques for robust watermarking," in *Proc. 4th Int. Workshop on Digital Watermarking*, Siena, Italy, Sep. 15–17, 2005, vol. 3710, Lecture Notes Comput. Sci.,pp. 418–431.
- [16] C. Schmid, R. Mohr, and C. Bauckhage, "Evaluation of interest point detectors," *Int. J. Comput. Vis.*, vol. 37, no. 2, pp. 151–172, 200
- [17] K. Mikolajczyk and C. Schmid, "Scale & affine invariant interest point detectors," *Int. J. Comput. Vis.*, vol. 60, no. 1, pp. 63–86, 2004.
- [18] T. Lindeberg, "Features detection with automatic scale selection," *Int.J. Comput. Vis.*, vol. 30, no. 2, pp. 79–116, 1998.
- [19] 19 Li-Li, M. Zhang, and Z. Pan, "Image watermarking algorithm resilient to geometrical transform," J. Zhe-Jiang Univ. (Eng. Sci.), vol. 38, no.2, pp. 141–144, 2004.