ICICS-PCM 2003
15-18 December 2003
Singapore

Image Watermarking System Based on Centroid Projection

Shih-Wei Sun and Pao-Chi Chang

Electrical Engineering, National Central University, Chung-Li, Taiwan
e-mail: {swsun, pcchang}@vaplab.ee.ncu.edu.tw

Abstract

We propose an image watermarking system aiming at reducing false positive rate based on the location changes of centroids of several selected points. Two secret keys are used to determine the number of selected points and the random number seed in the LL band of DWT domain, respectively, of the original and watermarked images. The number of modified pixel values is basically a tradeoff between the transparency of the image quality and the robustness against attacks. The two secret keys and the calculated centroid should be sent as the side information along with the watermarked image from the embedding end to the extracting end. The experimental results show that the proposed system yields not only high robustness against filtering and compression attacks but also low false alarm rates after attacking.

1.Introduction

The digital watermarking technique has been developed for several years. There were several approaches designed for embedding secret unique information into multimedia, such as image, audio, or video. In addition to robustness, low false alarm rate is also an important issue. False alarm includes two cases: the first one is the false positive, which means a watermark is detected but there is no watermark embedded. On the other hand, in a false negative case, no watermark is detected but a watermark is actually embedded in media. In this work, we concern not only the robustness, but also the false alarm problems, in particular, the false positive problem. The Discrete Wavelet Transform (DWT) and secret key based watermarking techniques, which were proposed in the past, yield high robustness and therefore serve as the starting points of the proposed system.

The multiresolution watermarking for digital images was proposed by Hsu and Wu [1]. A binary watermark is embedded into a gray-level image, both in DWT domain [2]. They considered that the components in the lowest frequency band of the image are left unmodified, because the modification for the coefficients here will affect the image quality a lot. In [3], Xei and Arce proposed a class of authentication digital watermarks for secure multimedia communication to consider the rank-order relationship in local areas throughout the lowest level of the DWT. They adopted SPIHT in their embedding algorithm in considering the parent-child relationship. Recently, the DWT based watermark system were also proposed by Oh, Park, and Kim in [4], and they considered that the method in DWT domain is robust to common signal distortions, noticing the trade-off between the quality, capacity, and robustness in designing of their system. From the above works, the signals in DWT domain provide high robustness against several attacks, especially the low frequency ones. Therefore, in our proposed method, the signals in that frequency band will be most concerned.

The secret and public key image watermarking schemes for image authentication and ownership verification was proposed by Wong and Memon in [5]. In their method, if the key is correct, the system will generate a proper watermark. If the key is incorrect, or if the image was not watermarked, the watermark extraction algorithm will return an image that resembles random noise. The aim of our proposed method is quite similar to their approach, and their system can detect if the image is watermarked or not, while the extractor owns the correct key.

In [6] [7], Low and Maxemchuk proposed the document identification for copyright protection using centroid detection and performance comparison of two text marking methods. They modify the position of the text words for a few pixels and calculate the centroid of profile in vertical and horizontal directions in both embedding and extracting end to identify if the document belongs to them or not. In their approach, they were dealing with text files, so the processing media were binary images. By shifting the position of the words or lines for a few pixels, the centroid can be removed. In our proposed method, we modify the pixel value instead of shifting pixels. In addition, the media that we process are real images, not binary text files, so the situation is quite different from their approach.

In this paper, we propose an image watermarking system based on centroid projection using two secret key with two centroid as the side information. The secret information are embedded in the LL band of DWT domain to provide the high robustness. The number of modified coefficients should be less enough to keep the transparency requirement, and the modification value should be adjusted toward certain direction moving the centroid that can be distinguished from the un-watermarked image, providing low false positive and false negative rates.

2.Centroid Calculation

For a given set $R \subset \{(x_1,y_1),(x_2,y_2),\cdots,(x_N,y_N)\}$ of random selected points in the LL band of DWT domain based on the two secret keys decided by the users, the centroid is defined as: $cent(R) = E(X|X \in R)$, where X contains the information about x and y directions. Hence,

$cent(R)$ should be thought as the position mean of the selected points.

In our proposed method, the original centroid (x_o, y_o) can be calculated from R in the un-watermarked image. When the coefficients are modified by the watermarking system, the centroid should be calculated again from R to generate (x_w, y_w). If the modifying factor is decided properly, the centroid will be shifted for a small distance from (x_o, y_o) and (x_w, y_w) for the un-watermarked and watermarked images, respectively. In Fig. 1, the concept is depicted as a two dimensional graph.

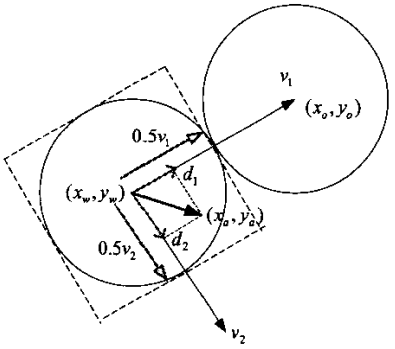


Fig. 1 Centroid calculation

(x_o, y_o) and (x_w, y_w) are the original and watermarked centroids. In the extracting end, the centroid can be calculated from the two secret keys, which is identical to the one in the embedding end. At this time, the vector v_1 can be formed by the two points (x_o, y_o) and (x_w, y_w) on the basis of x- and y- axis. Meanwhile, the vector v_2, which is orthogonal to vector v_1, can be also obtained.

Because the selected points are all in the LL band of DWT domain, even if the watermarked image is attacked by others, the coefficients will not be modified too much in order to keep the quality of the image. Therefore, the centroid calculated from the points should not be moved too much, such as within $\pm 0.5v_1$ and $\pm 0.5v_2$ in the two basis vectors. The moving vector will be the combination of the two basis v_1 and v_2. In other words, the centroid movement will be limited in the unit circle based on radius of $0.5 \cdot \|v_1\|$ and the center is (x_w, y_w). On the other hand, the centroid of the attacked image not watermarked will be limited in the unit circle based on radius $0.5 \cdot \|v_1\|$ and the center is (x_o, y_o).

We assume that the attacked centroid is Gaussian distributed, and the density function can be written as:

$$f(x,y) = A \exp\left\{-\frac{1}{2(1-r^2)}\left(\frac{(x-\eta_1)^2}{\sigma_1^2} - 2r\frac{(x-\eta_1)(y-\eta_2)}{\sigma_1\sigma_2} + \frac{(y-\eta_2)^2}{\sigma_2^2}\right)\right\}$$
$$\text{where } A = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}},\ |r|<1 \qquad (1)$$

The mean value of the watermarked centroid is $(\eta_1, \eta_2) = (x_w, y_w)$, and the mean of the original image is $(\eta_1, \eta_2) = (x_o, y_o)$. The variances and correlation coefficients can be calculated by a lot of experiments about attacks for the watermarked image and un-watermarked images.

So far, the general concept about centroid calculation proposed in our method is realized. The watermark embedding and extracting process will be illustrated in detail in the next section, and the other items of Fig. 1 will also be used for discussion in the section.

3.System Architecture

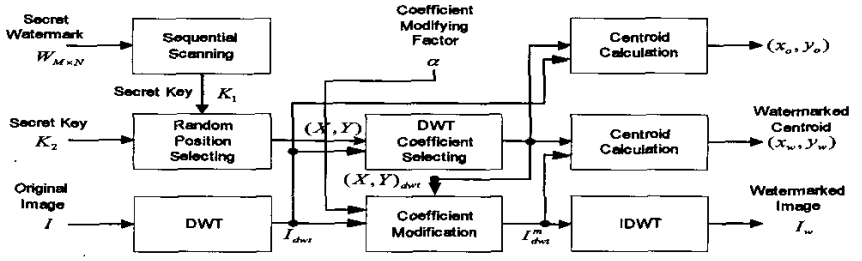
From the concept described in the previous section, the centroid calculation can be utilized in our proposed method and playing an important role both in embedding and extracting end.

In this paper, we propose an image watermarking system based on centroid projection. The system can be applied to the whole existing digital images to provide robustness, security, and low false alarm rate. The block diagram of the system is depicted in Fig. 2.

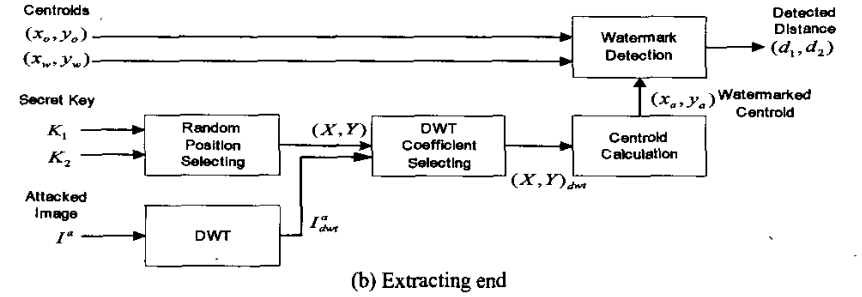
3.1. Embedding

Fig. 2. (a) shows the watermark embedding process. The watermark $W_{M\times N}$ is embedded into the original image I by the two secret keys K_1 and K_2. Assuming the watermark $W_{M\times N}$ is well defined and kept by the users, and it will be a binary two-dimensional matrix. Therefore, by the conventional sequence scanning from up to bottom, from left hand side to right, the secret key K_1 can be formed. The secret key K_2 is also used to determine the random selected positions. Basically speaking, K_2 is utilized to decide the number of modified positions of the LL band of the DWT domain. After random position selecting, (X,Y) can be decided, where (X,Y) represents the position set of the selected ones.

The original image I is Discrete Wavelet Transformed (DWT) into wavelet domain to form I_{dwt}. At this time, the LL band signals of DWT domain are the most important that we concerned, because the coefficients in LL band are the candidates that we are willing to process. By considering the random selected positions (X,Y) and the LL band of the original wavelet domain signal I_{dwt}, the positions in DWT domain $(X,Y)_{dwt}$ can be determined. The selected positions of $(X,Y)_{dwt}$ should be sent to three parts of the block diagram to modify the coefficients and calculate the centroids before and after coefficients modification. Therefore, the modified DWT signals I^m_{dwt}, original centroid (x_o, y_o), and watermarked centroid (x_w, y_w) can be obtained, respectively. Finally, after IDWT, the watermarked I_w is completely made.



(a) Embedding end



(b) Extracting end

Fig. 2 The block diagrams: (a) the embedding end (b) the extracting end

The coefficient modification is described as follows:

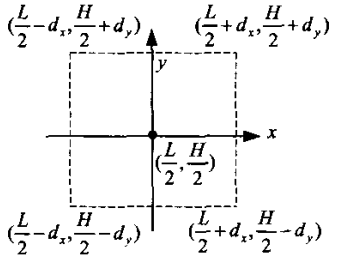


Fig. 3 The geometric relationship of the selected points in DWT domain

Fig. 3 represents the geometric relationship among selected points in DWT domain, where the LL band of DWT domain is $L \times H$. If one point is selected, meanwhile, the other three points are selected. This is an example of $K_2 = 1$. If the selected point is (x_s, y_s), the difference vector $(d_x, d_y) = (\frac{L}{2} - x_s, \frac{H}{2} - y_s)$. Therefore, the geometric relationship can be derived as in Fig. 3.

Moreover, the centroid should be moved to other place from (x_o, y_o), so the design target here is to move the centroid to one direction. The following equation provides an example:

$$\begin{cases} I^m_{dwt}(\frac{L}{2}-d_x, \frac{H}{2}+d_y) = I_{dwt}(\frac{L}{2}-d_x, \frac{H}{2}+d_y) + \alpha \\ I^m_{dwt}(\frac{L}{2}+d_x, \frac{H}{2}+d_y) = I_{dwt}(\frac{L}{2}+d_x, \frac{H}{2}+d_y) + \alpha \\ I^m_{dwt}(\frac{L}{2}-d_x, \frac{H}{2}-d_y) = I_{dwt}(\frac{L}{2}-d_x, \frac{H}{2}-d_y) - \alpha \\ I^m_{dwt}(\frac{L}{2}+d_x, \frac{H}{2}-d_y) = I_{dwt}(\frac{L}{2}+d_x, \frac{H}{2}-d_y) - \alpha \end{cases} \quad (2)$$

where α is the coefficient modifying factor. In the example, the centroid will move in y-direction. The modifying factor and sign of the modifying relation can be adjusted by the designers.

In the embedding end, the secret key K_1, K_2, and original centroid (x_o, y_o), watermarked centroid (x_w, y_w) should be sent as the side information to the extracting end.

3.2. Extracting

Fig 2. (b) shows the watermark extracting process. The attacked image I^a should be first DWT as in the embedding end to obtain the LL band signals in wavelet domain I^a_{dwt}. Using the same secret key K_1, K_2, the random positions (X,Y) can be selected the same as in the embedding end. Through the centroid calculation, the watermarked centroid (x_a, y_a) can be obtained from the position in the DWT domain $(X,Y)_{dwt}$. At this moment, the three centroids are all obtained: (x_o, y_o), (x_w, y_w) from side information, and (x_a, y_a) from the attacked image.

By comparing the three centroids, the attacked image can be decided if the image is watermarked or not. If the detected value can satisfy both the criteria:

$$\begin{cases} -0.5v_1 \le d_1 \le 0.5v_1 \\ -0.5v_2 \le d_2 \le 0.5v_2 \end{cases}, \quad (3)$$

we can decide that the image is watermarked; otherwise, the image is not watermarked. The vectors are depicted in Fig. 1. The conditions are listed below. (x_a, y_a) is the detected centroid, v_1 is the difference vector between (x_o, y_o) and (x_w, y_w), $v_2 \perp v_1$ and $\|v_2\| = \|v_1\|$, and $d_1 = \frac{(x_a, y_a)\cdot v_1}{v_1 \cdot v_1} v_1$, $d_2 = \frac{(x_a, y_a)\cdot v_2}{v_2 \cdot v_2} v_2$. The watermark detection region is enclosed by the outer dash lined rectangle. If the detector have the correct key, but the image is not watermarked, the detected centroid (x_a, y_a) will be existed near the unit circle center by (x_o, y_o). However, the divergence degree is not the most that we concerned, because the aim is to decide if the image is watermarked or not. The most probable event that happened in detecting process is "not watermarked", if the key is not correct or the image is really not watermarked.

On the other hand, if the detector does not have the correct key, the centroid will not exist in the small area of the watermarked region, and the detected distance will be much larger than the criteria.

4. Experimental Results

In the experiments, the false positive, false negative analyses are tested, and the Checkmark [8] evaluation is also performed to show the robustness.

In the simulation, the most wildly used test image Barbara, Lena, and Pepper are selected. The image size is 512x512, and DWT is two level with 9-7 filter. The PSNR of the three images are all 69.69 dB after watermark embedding. The false alarm rate analysis is according to [9].

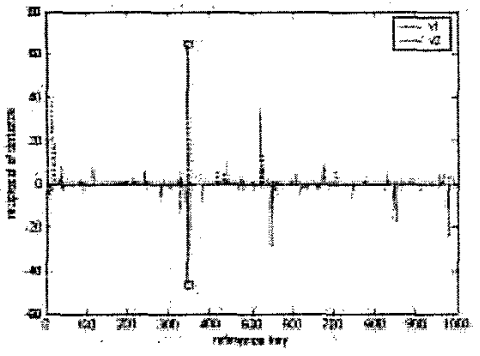


Fig. 4 The false positive analysis

The false positive analysis is depicted in Fig. 4. The correct $K_1 = 2$, and $K_2 = 347$. The horizon axis is the number of reference key generated by randomly selected, and the vertical axis represent the reciprocal of detected distance. If the detected value can satisfy the criteria in (3), the image is decided as watermark detected. On the other hand, if the reciprocal of detected distance v_1, v_2 can be both larger than 2, the watermark is detected. When $K_1 = 2$, and $K_2 = 347$, the reciprocal of detected distance v_1, v_2 are both larger than 40, so the watermark is detected successfully.

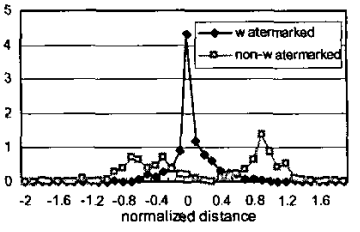


Fig. 5 The false negative analysis

The false negative analysis is depicted in Fig. 5. The norm of d_1 and d_2 are both tested. If the image is watermarked, the detected distance is probably near 0, if the image is not watermarked, the detected distance is probably near 1, and the result shows that the decision boundary could be decided as 0.5.

Table 1 The watermark detection –watermarked images

attack	wateramrk detection	barbara	lena	pepper
Wiener filter 3x3	Y, Y, Y	(0.1,-0.1)	(0.1,-0.0)	(0.2,-0.1)
Soft thresh 5x5	Y, Y, Y	(0.0,-0.0)	(-0.1,0.3)	(0.1,0.2)
Median filter 3x3	Y, Y, Y	(0.4,-0.3)	(0.0,0.1)	(0.3,-0.2)
Gaussian 5x5	Y, Y, Y	(-0.0,0.0)	(-0.1,0.1)	(0.0,0.0)
JPEG Q=90	Y, Y, Y	(0.0,-0.0)	(-0.1,0.1)	(-0.0,0.0)
JPEG Q=30	Y, Y, Y	(-0.1,0.1)	(0.3,-0.2)	(0.1,-0.3)
JPEG Q=25	Y, N, Y	(0.1,-0.1)	(0.7,-0.7)	(0.0,0.2)
JPEG-2k bpp=8.0	Y, Y, Y	(-0.0,0.0)	(0.0,-0.0)	(0.0,-0.0)
JPEG-2k bpp=0.4	Y, N, Y	(0.1,-0.1)	(0.5,-0.6)	(-0.1,0.4)

Table 2 The watermark detection –non-watermarked images

attack	wateramrk detection	barbara	lena	pepper
Wiener filter 3x3	N, N, N	(1.0,-0.8)	(1.0,-0.7)	(1.1,-0.4)
Soft thresh 5x5	N, N, N	(0.9,-0.6)	(0.8,-0.4)	(0.9,-0.2)
Median filter 3x3	N, N, N	(1.3,-1.0)	(0.9,-0.5)	(1.2,-0.6)
Gaussian 5x5	N, N, N	(0.9,-0.7)	(0.9,-0.6)	(1.0,-0.3)
JPEG Q=90	N, N, N	(1.0,-0.8)	(1.0,-0.7)	(1.0,-0.4)
JPEG Q=30	N, N, N	(0.8,-0.6)	(1.2,-0.8)	(1.0,-0.6)
JPEG Q=25	N, N, N	(1.2,-0.9)	(1.2,-1.0)	(0.9,-0.1)
JPEG-2k bpp=8.0	N, N, N	(1.0,-0.7)	(1.0,-0.7)	(1.0,-0.4)
JPEG-2k bpp=0.4	N, N, N	(1.0,-0.8)	(1.5,-1.3)	(0.6,0.2)

Table 1 shows the Checkmark attacked detection values, non-geometric attacks. The Wiener, soft thresh, Median,

and Gaussian filters with 3x3 or 5x5 sizes are selected to shown. The results show that the watermark can all be detected under the filtering attacks. As we mention to the compression attacks, DCT-based and DWT-based compression standards of JPEG and JPEG-2000 are selected for experiments. In high bit-rates, the watermark can be detected, but if the low-bit rate compression attacks on the watermarked images, the detection result will be relatively weak, because of the image quality strong degrading. Table 2 shows the Checkmark attacked detection values, the situation is the same as in Table 1, except the images are not watermarked. In the table, the results show that if the images are not watermarked, the watermark will not be detected as absence. Not only the filtering attacks, but also the compression attacks with DCT and DWT-based, the detection result is perfect.

The results show that the watermark can still survive in most of attack environments. As we mention to the geometric attacks, the attacked images can be recovered by the method proposed in [10].

5. Conclusions

In our proposed method, we have developed an image watermarking system based on centroid projection, which provides low false alarm rate and is robust to almost all difficult environments attacked by Checkmark. The experimental results show that the characteristic mentioned above is existed.

References

[1] C. T. Hsu and J.L. Wu, "Multiresolution watermarking for digital images," IEEE Trans. Circuits and Systems II: Analog and Digital Signal Processing, Vol. 45, No. 8, pp. 1097-1101, Aug. 1998.

[2] I. Daubechis, "Ten lectures on Wavelets," CBMS-NSF regional conference in applied mathematics: 61,1992.

[3] L. Xie and G.R. Arce, "A class of authentication digital watermarks for secure multimedia communication," IEEE Trans. Image Processing, Vol. 10, No. 11, pp. 1754-1764, Nov. 2001.

[4] S.H. Oh, S.W. Park, and B. J. Kim, "DWT based watermark system," IEEE International Conference on Consumer Electronics, Digest of Technical Papers, pp. 192-193, 2002

[5] 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.

[6] S. H. Low, N. F. Maxemchuk, and A. M Lapone, "Document identification for copyright protection using centroid detection," IEEE Trans. Communications, Vol. 46, No. 3, pp. 372-383, Mar. 1998.

[7] S. H. Low and N. F. Maxemchuk, "Performance comparison of two text marking methods," IEEE Trans. Selected Areas in Communications, Vol. 16, No. 4, pp. 561-572, May. 1998.

[8] http://watermarking.unige.ch/Checkmark/

[9] I.J. Cox, M.L. Miller, and J.A. Bloom, "Digital watermarking", Morgan Kaufmann Publishers, 1/e, USA, 2002

[10] S.W. Sun, T.T. Lu, and P.C. Chang, "Image Watermarking Synchronization by Significant MSB Plane Matching," IEEE Pacific Rim Conference on Multimedia, pp. 468-476, Hsin-Chu, Taiwan, Dec. 2002