A DIGITAL WATERMARKING SCHEME BASED ON ICA DETECTION

Ju Liu[1, 2], *Xingang Zhang*[1], *Jiande Sun*[1] and *Miguel Angel Lagunas*[2, 3]

[1]School of Information Science and Engineering, Shandong University, Jinan 250100, China

[2]Telecommunications Technological Center of Catalonia, Barcelona 08034, Spain

[3]Dept. of Signal Theory & Communications, Universitat Politècnica de Catalunya, Barcelona 08034, Spain

Email: [1]zhangxg1978@sina.com, [2]juliu@sdu.edu.cn, [3]m.a.lagunas@cttc.es

ABSTRACT

The scheme proposed in this paper combines independent components analysis (ICA) with discrete wavelet transform (DWT) and discrete cosine transform (DCT). Firstly, the original image is decomposed by 2-D DWT and the detail sub-bands are reserved. Then, the approximate image is transformed by DCT and embedded with watermark. The watermark is detected through ICA. The simulation results demonstrate its good performance of the robustness and invisibility. The watermark detection is improved greatly compared to the traditional subtracting detection scheme.

1. INTRODUCTION

With the prevalence of the multimedia network, the transmission and access of digital medias (such as digital images, digital video, digital audio, and so on) become more and more convenient. Though it helps people to share the information, at the same time it brings new problem to the authentication of copyright. Digital watermark is developed to solve this problem. Digital watermark is the digital signals embedded into the digital products. It may be any signal that can be used as signs or marks, such as images, letters, signs, numbers and so on. Among the research of watermark, the robust watermark used for the protection of the copyright is discussed mostly, the most important properties of which are the robustness and the imperceptibility.

The method of how to embed digital watermark can be classified into two kinds: methods in temporal or spatial domain and those in transform domain. The anterior method is generally to change the brightness or chroma values of an image directly, but its robustness is weak. The method of Least Significant Bits is a representative one and one of the earliest methods of watermark [1]. It replaces the LSB of the original image with the bits of watermark or the logical result of the two images. The watermark embedded by this method is easy to be got by others and isn't very robust against image processing. In order to improve the capability, there are many improved methods [2], such as using pseudo-random sequence, modifying the LSB randomly, and you can get the correct embedded sequence only with the key. The methods performed in transform domain use a certain arithmetic transform to represent the original image in another domain, and hide the watermark by changing the transform coefficients, and then transform inversely to gain the watermarked image. Today, this kind of method usually uses wavelet analysis, transforms in frequency domain, DCT, etc. Cox *et al.* embed the watermark into the low-frequency AC coefficients in the DCT domain [3]. This scheme is more robust against JPEG compression, noise etc. compared with the methods in temporal or spatial domain. With the research being carried on, adaptive schemes combined with the human vision system (HVS) emerge [4][5]. These schemes make use of HVS, such as the JND in the DCT domain, exploit the potential of information embedding to the utmost extent, and improve the robustness against attacks.

In this paper we also use HVS to improve the watermarking performance. The research on human vision shows the high-frequency edge details are the important factors affect human eyes [6], so we should keep their integrity. Accordingly we propose a embedding scheme combines DWT with DCT. First the original image is decomposed by 2-D DWT and the detail sub-bands are reserved. Then, the approximation image is transformed by DCT and embedded with watermark. The experimental results demonstrate its good unification between the robustness and invisibility.

Independent Components Analysis (ICA) is a novel signal processing and data analysis method developed in the research of blind signals separation. Using ICA, even without any information of the source signals and the coefficients of transmission channel, people can recover or extract the source signals only from the observations according the stochastic property of the input signals. It has been one of the most important methods of blind source separation and received increasing attentions in pattern recognition, data compression, image analysis and so on. In this paper ICA is used in the process of watermark detection. The watermarked and the original images are the observation and the source signals respectively, so the extraction of the watermark can be looked as the blind source separation. The simulation experiment shows the watermark can be better detected by using ICA.

2. STABILITY OF DWT AND DCT

2.1 Wavelet and Multi-resolution Analysis (MRA)

Wavelet analysis is useful signal processing methods especial in image processing. The main theory in Wavelet analysis is MRA that analyzes a signal in frequency domain in detail. It is compatible with perception of human eyes. Through wavelet analysis, an original image can be decomposed into an approximate image LL and three detail images LH, HL and HH. Using wavelet analysis on the approximate image again, four lower-resolution sub-band images will be got, and among them, the approximate image hold most of the information of the original image, while the others contain some details such as the edge. These details can be affected easily by the noise, some common image processing, etc, so they are not stable enough to hide information in them. Compared with them, the approximate image has its own superiority [6]. The table 1 gives the affection on each sub-image after certain image processing. It is presented by the normalized variance, and formula (1) tells us how to calculate the normalized variance.

$$Std = \sum_i \sum_j \frac{1}{N \times N} \cdot \frac{[I(i,j) - I'(i,j)]^2}{(Max(I(i,j)))^2} \qquad (1)$$

In formula (1), $I(i,j)$ and $I'(i,j)$ are the coefficients of sub-image before and after the image processing, respectively. The size of sub-image is $N \times N$.

Table 1: normalized variance of some wavelet sub-images at different levels of " LENA" after degradation ($\times 10^{-3}$)

	LL_2	LH_2	HH_2	LH_1	HH_1
Gaussian noise	2.6	17.0	17.8	23	34.3
3×3 media filter	0.2	1.1	1.7	1.5	2.3
JPEG (Cr=10.8)	0.2	1.0	1.8	1.6	2.9

Table 1 shows that after the degradation the normalized variance of approximate image is much smaller than any of the other detail image and this means it is more stable against outside attack. According to this, we should embed the watermark into the approximate image. On the one hand, we can gain the robustness, and on the other hand, we can keep the integrity of the details, and hence improve the imperceptibility.

2.2 DCT

The approximate sub-image in wavelet domain is difficult to be affected, so it is used for watermark embedding. But they are full of low-frequency ingredient of the original image, which is sensitive to human eyes, so the watermark embedded by modulating the coefficients will decrease the quality of the image [6]. According to this, DCT transform is used to do further decomposition to remove the correlation existing between rows and columns. And then the watermark is embedded in DCT domain. Thus the energy of watermark is distributed throughout the whole sub-image and a better invisibility can be obtained.

3. INDEPENDENT COMPONENTS ANALYSIS

Recently, the research on ICA (Independent Component Analysis) and its application has been the focus in the field of signal processing. To do ICA is to find a certain linear transform which can decompose the objective vectors and make the components of it as independent as possible. Though ICA comes from blind source separation or blind signal separation, it is used widely in many other fields, such as feature extraction, data compression and image analysis, etc.

3.1 Problem Formulation

Assume that there are *m* sensors and *n* source signals in the mixture system. The relationship between sources and observations is:

$$\mathbf{X}(t) = \mathbf{AS}(t) \tag{2}$$

Where the observations $\mathbf{X}(t)=[\mathbf{x}_1(t),\mathbf{x}_2(t),\cdots,\mathbf{x}_m(t)]^T$ are the simultaneous mixtures of *n* unknown sources $\mathbf{S}(t)=[\mathbf{s}_1(t),\mathbf{s}_2(t),\cdots,\mathbf{s}_n(t)]^T$. $\mathbf{s}_i(t), i = 1,2,\cdots,n$ are *n* mutually independent stochastic signals. The mixing matrix is $\mathbf{A}=(a_{ij}), i=1,2,\cdots,m, j=1,2,\cdots,n$.

By estimate matrix **W** we can recover the source **S** (t) by

$$\mathbf{Y}(t) = \mathbf{WX}(t) = \hat{\mathbf{S}}(t) \tag{3}$$

Many ICA algorithms have been proposed [7-9]. Now we give a simple description of the algorithm based on minimization of mutual information [7, 9] which we used to detect the watermark in this paper.

3.2 ICA Algorithm Based on Minimization of Mutual Information

Assuming that the sources are mutually statistical independent, we use the average mutual information between output $\mathbf{Y}(t)=[\mathbf{y}_1(t),\mathbf{y}_2(t),\cdots,\mathbf{y}_n(t)]^T$ and its components as the measure of statistical independence:

$$J(\mathbf{W}) = MI(\mathbf{Y};\mathbf{y}_1,\mathbf{y}_2,\ldots,\mathbf{y}_n) = \sum_{i=1}^{n} H(\mathbf{y}_i) - H(\mathbf{Y}) \tag{4}$$

in which $H(\mathbf{Y})$ and $H(\mathbf{y}_i)$ are joint entropy and marginal entropy, respectively. Without knowing the source signals and their probability, the formula (4) cannot be utilized directly. Here, we expand the entropy by Edgeworth expansion and truncate it into the fourth order cumulants [8], [9]. The main advantage of using the Edgeworth expansion lies in the ordering of terms according to their decreasing significance. Though it is formally equivalent to the Gram-Charlie expansion from a practical point of view, however, it represents a rearrangement of that series to one that has better asymptotic convergence properties.

The ICA learning algorithm based on the Edgeworth expansion can be given by:

$$\mathbf{W}(t+1) = \mathbf{W}(t) - \eta(t)(I - f(\mathbf{Y})\mathbf{Y}^T)\mathbf{W}(t) \tag{5}$$

where the $\mathbf{I}$ is an identify matrix, $\eta(t)$ is the step size and the $f(\mathbf{y})$ is a nonlinear function obtained by Edgeworth expansion of the entropy [8]. The details of the algorithm can be found in references [7-9].

4. WATERMARK EMBEDDING AND DETECTION

4.1 Watermark Embedding

First, the original image is decomposed into an approximate image LL_N and some detail sub-bands $LH_n, HL_n, HH_n (n = 1,2,...,N)$ by N level DWT. Then the approximate image LL_N is transformed by DCT, and we get the DCT coefficients matrix $F(u,v)$. The watermark $M(i)(i = 1,2,...,I)$ is a sequence chosen according to $N(0,1)$ (where $N(\mu,\sigma^2)$ denotes a normal distribution with mean μ and variance σ^2), and I is the length of watermark sequence. The watermark is embedded into the low-frequency AC coefficients of $F(u,v)$ by using the following formula

$$F'_{k+i} = F_{k+i} + a \times M(i) \quad i = 1,2,......,I \tag{6}$$

where F_k is the K^{th} coefficient after the matrix $F(u,v)$ is zigzag scanned and *k* is the jumping-off point of watermark embedding. a is the embedding intensity factor.

The modulated DCT coefficients are inverse zigzag scanned into 2-D DCT coefficients matrix $F'(u,v)$ which then is transformed in to the approximate image LL'_N by 2-D inverse DCT. At the last, LL'_N and $LH_n, HL_n, HH_n (n = 1,2,...,N)$ are combined and transformed into the watermarked image by 2-D inverse DWT of the same level.

4.2 Watermark Detection by using ICA

The ICA method is used in the watermark detection. The process of detection is as following.

Firstly, the image which needs to be detected is transformed by N level DWT, and the approximate image LL''_N we get is also transformed by 2-D DCT and then zigzag scanned into 1-D sequence. According to the position the watermark embedded, we choose the corresponding DCT coefficients sequence $F'' = \{F''_{k+1}, F''_{k+2}, ..., F''_{k+I}\}$ as one observation in performing ICA, and the DCT coefficients in the same position is picked out as another observation. Formula (5) is used to train the de-mixing matrix **W**, and then we use formula (3) to separate and we can obtain the watermark sequence M'.

In order to justify the watermark sequence, measure the similarity of M' and M by

$$p = \sum_i M(i) \cdot M'(i) / sqrt(\sum_i (M'(i)\ )^2) \tag{7}$$

if $p > T_0$, where T_0 is a threshold, we consider that there is a watermark in the detected image, otherwise no. The threshold T_0 is five times of the variance of $N(0,1)$, so T_0=5 is appropriate.

5. SIMULATION RESULTS USING THE PROPOSED APPROACH

In this experiment, we transformed the original image by 2-level DWT. Randomly generate 600 watermark sequences whose length I are all 1024, and embed the 300th into the original image. The embedding factor a=0.06, and the jumping-off point of DCT coefficient is $K = 30$. The quality of watermarked image is evaluated objectively by

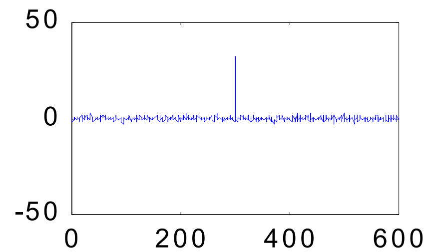
$$PSNR = 20\lg \frac{255}{\sqrt{\frac{1}{m \times n}\sum_i \sum_j (I(i,j) - I'(i,j))^2}} \quad (8)$$

Figure 1a is the original image LENA (256×256×8bit). Figure 1b is the watermarked image, *PSNR*=42.49. From this *PSNR*, we can see that the objective quality of the watermarked image is very good. Subjectively the watermarked is the same as the original. The invisibility is performed. Figure 1c is the detection result, we can find the corresponding value is outstanding ($p = 31.8$).

In order to test the robustness of the proposed scheme, we attack the watermarked image in several ways. Attack 1 is performing JPEG compression with quality=5, compression rate Cr=31 and *PSNR*=24.31; Attack 2, add Gaussian noise with mean=0, variance=0.09 and *PSNR*=11.89; Attack 3, median filter with 7×7 filter and *PSNR*=24.35; Attack 4, crop the top 3 lines and bottom 3 lines of the watermarked image, *PSNR*=24.74; Attack 5, rotate the image 1.2°, *PSNR*=24.73; Attack 6, resize the image by 25% and then amplify to the original size, *PSNR*=23.28. The experimental results show (in Figure 2) that the image is degraded greatly after each attack, but the detection values are still higher than the threshold T_0, so the watermark can be detected correctly. Figure 2a-c are some detection results. Moreover, if multiple watermarks are embedded, each of them can be detected correspondingly. Fig. 2(d) is the response of detector with two watermarks. Because the scheme is combined with DWT and the watermark is embedded into the low-frequency part, the robustness is very good against JPEG2000 compression which is also based on DWT. After the image is compressed with JPEG2000 (compression Cr=49, *PSNR*=24.73), the response of detector is still as high as 13.94.

Table 2 is the comparison of our scheme and others proposed in [3], [4]. The result is obtained under the condition that the watermarks embedded into image LENA are in the same length and the watermarked images are of the same *PSNR*. Through comparison we can see that the proposed scheme is more robust than the others.

(a) Original image (b)Watermarked



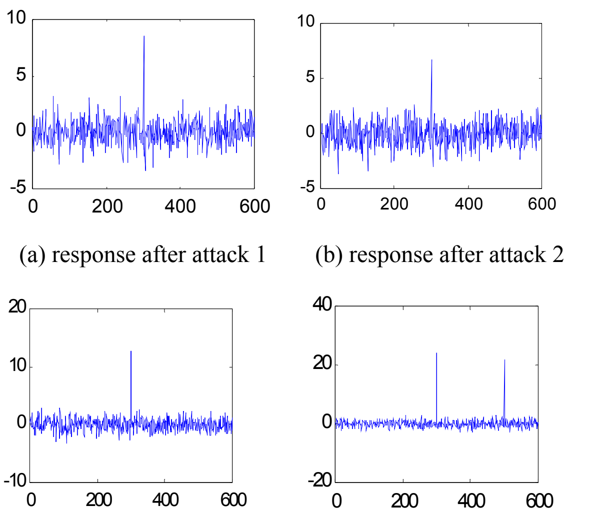
(c) Detection result

Fig.1 original and watermarked LENA image,

Though we can also pick out the watermark if the DCT coefficients of the original image $F = \{F_{k+1}, F_{k+2}, \ldots, F_{k+l}\}$ and those of the watermarked image $F'' = \{F''_{k+1}, F''_{k+2}, \ldots, F''_{k+l}\}$ are subtracted directly, the result is worse than that using ICA. Table 3 is the comparison of detection results using ICA method or direct subtraction against Gaussian noise attack. From the table we can see the ICA method can still get reliable result when the direct subtraction method fails. Considering other attacks our ICA method is also better than direct subtraction method to some extent. So it is clear that the use of ICA method improves the accuracy of watermark detection greatly.

6. CONCLUSION

In this paper, according to the human visual masking model, our watermark embedding scheme analyzes the stability of wavelet coefficients of the image, combines the multi-resolution wavelet decomposition with discrete cosine transform and through modulating the middle- and low-frequency coefficients matches its robustness and invisibility. The simulation experiments show the robustness has been improved and at the same the high image quality is kept. Besides, the theory of blind separation is used in the watermark field and a novel ICA method makes great improvement in watermark detection and makes this scheme superior.



(a) response after attack 1 (b) response after attack 2

(c) response after attack 6 (d) response of two watermarks

Fig. 2 detection results

Table2: robustness comparison of some watermark scheme

Attack	1	2	3	4	5	6
Our scheme	8.62	6.74	6.6	5.94	5.76	12.65
Ref. [3]	5.03	3.01	4.06	3.9	3.1	5.79
Ref. [4]	9.96	3.11	6.07	5.73	3.86	11.09

Table 3: Detection results using ICA or direct subtraction

Variance of Gaussian noise		0.06	0.07	0.08	0.09
a=0.06 K=35	ICA	7.03	6.53	6.13	5.85
	Direct subtraction	5.40	4.77	4.26	3.81
a=0.047 K=30	ICA	6.47	6.11	5.86	5.54
	Direct subtraction	4.84	4.35	3.99	3.66

ACKNOWLEDGEMENT

This project is partially supported by The National Natural Foundation of China (No. 30000041), The Natural Foundation of Shandong (No. Y2000G12 and 13) and The Excellent Young Scientist Award Foundation of Shandong (No. 01BS04). Ju Liu is granted by the Ministry of Education and Culture of Spain (No. SB2000-0212).

7. REFERENCES

[1] R. G. van Schyndel, A. Z. Tirkel and C. F. Osborne, " A Digital Watermark, " *Proc. Int. Conference in Image Processing,* vol.2, pp. 86-90, 1994.

[2] R. G. van Schyndel, A. Z. Tirkel and C. F. Osborne, " Towards a Robust Digital Watermark," *ACCV'95 Conference,* vol.2, pp.504-508, December 5-8, 1995.

[3] I. J. Cox, J. Kilian, T. Leighton and T. Shamoon, " Secure Spread Spectrum Watermarking for Multimedia, " *IEEE Trans.on Image Processing,* vol.6, no.12, pp.1673-1687, Dec, 1997.

[4] J. W. Huang, Q. S. Yun and W. D. Cheng, " Image Watermarking in DCT: An Embedding Strategy and Algorithm," *Acta Electronica Sinica,* vol.28, no.4, pp.57-60, 2000.

[5] C. I. Podilchuk, " .Image-adaptive Watermarking Using Visual Models, "*IEEE Journal on Selected Areas in Communications,* vol.16, no.4, pp.525-539,1998.

[6] Y. Zhou, Q. Ye and T. Xu, " A New Scheme of Image Watermarking Based on Wavelet and Cosine Transform," *Acta Electronica Sinica,* vol.29, No. 12, pp.1693-1695, 2001.

[7] S. Amari, A. Cichocki and H. H. Yang, " A New Learning Algorithm for Blind Signal Separation, " *Advances in Neural Information Processing Systems,* MIT Press, pp. 657-663, 1996.

[8] J. Liu, K. B. Nie, and Z. He, " Blind Separation by Redundancy Reduction in A Recurrent Neural Network, " *Chinese Journal of Electronics,* vol.10, no. 3, pp. 415-419, 2001.

[9] Z. He, J. Liu, and L. X. Yang, " Blind Separation of Images Using Edgeworth Expansion Based ICA Algorithm, " *Chinese Journal of Electronics,* vol.8, no.3, pp. 278-282, 1999.