2003 International Symposium on Advanced Intelligent Systems
September 25-28, 2003, Jeju, Korea

EXTRACTION OF WATERMARKS BASED ON INDEPENDENT COMPONENT ANALYSIS

Hien Duy Thai, Zensho Nakao, & Yen-Wei Chen

University of the Ryukyus, Okinawa 903-0213, Japan,
Email: {tdhien, nakao, chen}@augusta.eee.u-ryukyu.ac.jp

Abstract **- We propose a new logo watermark scheme for digital images which embed a watermark by modifying middle-frequency sub-bands of wavelet transform. Independent component analysis (ICA) is introduced to authenticate and copyright protect multimedia products by extracting the watermark. To exploit the Human visual system (HVS) and the robustness, a perceptual model is applied with a stochastic approach based on noise visibility function (NVF) for adaptive watermarking algorithm. Experimental results demonstrated that the watermark is perfectly extracted by ICA technique with excellent invisibility, robust against various image and digital processing operators, and almost all compression algorithms such as Jpeg, jpeg 2000, SPIHT, EZW, and principal components analysis (PCA) based compression.**

1. INTRODUCTION

A digital watermarking is a secretly hidden code irremovable, imperceptible, and robustly embedded in the host data. It typically contains information about the origin, ownership, status or any other transaction information pertaining to the host data. Most of the existing watermark techniques are based on some assumptions for the watermark detection and extraction, and some require the previous knowledge of the watermark, watermark location, the strength, the threshold or original image. In order to insert an invisible watermark, the current trend has been to model the Human Visual System (HVS) and specify as a perceptual masking function which yields the allowable distortion of any pixel. This complex function compiles contrast, luminance, color, texture and edges. The watermark is then inserted in the transform domain and the inverse transform is computed. The watermark is finally adjusted to satisfy the contrast on the pixel distortions [1] [3].

This paper present a new ICA watermarking scheme that does not rely on the original information such as original image, the watermark, the strength...etc for recovering the watermark. Watermark is a signature $\{+1,-1\}$ text image, permuted to become pseudo-random sequence, and embedded in the DWT domain. We added those pseudo-random sequence watermarks to the DWT coefficients of the two detail I_l^{LH} , I_l^{HL} sub-bands of the host image, where l is final decomposition level. A robust watermarking technique should largely exploit the characteristics of the HVS for more effective watermark hiding, therefore in this proposed method, each watermark value is multiplied by a strength parameter and in combination with the mask function (NVF) which is computed based on a stochastic model for adaptive embedding [2]. For the watermark detection/extraction a new method based on ICA is introduced, which only uses the de-mix key, and one can extract the watermark.

Our experiment results show the validity of the technique both from the point of view of watermark invisibility and from the point of view of robustness against the most common attacks: the proposed system is robust to compression image such as Jpeg and Jpeg2000 [10], SPIHT [9], EZW [8] and PCA based compression, and demonstrated to be resistant to median filtering, low-pass filtering, adding noise, denoising, cropping and resizing.

2. THE EMBEDDED ALGORITHM

2.1. Scrambling of the Watermark

Our watermark is a digital text image signature $\{+1,-1\}$. By using such a watermark, it is more robust against attacks because the signatures can always preserve a certain degree of structural information, which are meaningful and recognizable, and the extracted watermark also can be easier to verify by human eyes rather than correlation method. Also to decrease the effect of the image cropping attacks, it prevents the watermark from tampering or unauthorized access by attackers. For the scrambling of the watermark, the pixels of the watermark are pseudo-randomly permuted to form a new watermark image. Let W and W_p be the original and permuted watermark, that is

$$W_p = \{ w_p(i,j) = w(i',j') \mid 0 < i,i' < M \text{ and } 0 < j,j' < N \}$$

where pixel at (i',j') is mapped to pixel at (i,j) in a pseudorandom order, M and N are sizes of the watermark. Fig.1 shows the binary text image of size 64x64 used for watermark and the corresponding permutation of the watermark.

(a) (b)

Fig.1. (a) Watermark, (b) Permutation of the watermark

2.2. Embedding watermark

The image to be watermarked is decomposed into a number of levels by wavelet transform (l = 1, 2, 3...). A watermark is inserted in the mid frequency sub-bands by modifying the wavelet coefficients belonging to the two detail bands at final level (I_l^{LH} , I_l^{HL}). This choice has been made based on an optimal compromise among robustness, invisibility and the attack. With this choice the watermark could be made more robust and effective against image degrading.

The watermark is permutated as a pseudorandom binary $\{+1,-1\}$ sequence, is inserted by modifying the wavelet coefficients belonging to I_l^{LH} and I_l^{HL} sub-bands at final level l.

The stochastic model of cover image is applied to content adaptive watermark by computing an NVF. We consider the locally i.i.d. (independent identically distributed) non-stationary Gaussian. Based on information of the host image, one can compute the NVF function [2]. Fig 2 shows the (a) permuted watermark, (b) NVF function of "Lena", masks at (c) texture and edge areas and (d) flat areas.

(a) W (b)NVF (c)(1-NVF).W (d)NVF.W

Fig.2. (a) watermark, (b) NVF masking function of "Lena", (c) watermark at edge and texture region, (d) watermark at flat regions

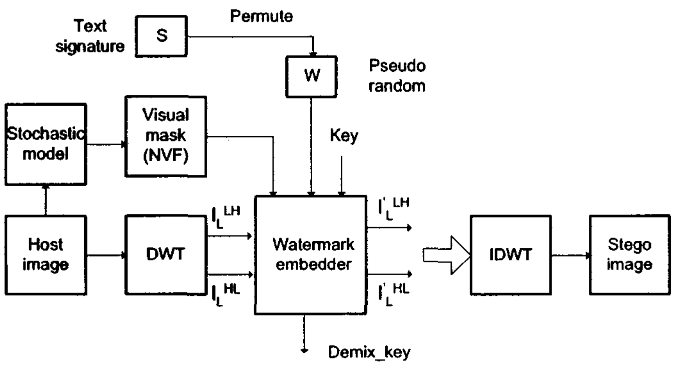


Fig.3. The proposed watermark embedding scheme

The watermark is embedded by the following formula:

$$\begin{aligned} I_l'^{LH}(i,j) &= I_l^{LH}(i,j) + A^{LH}.\alpha(1-NVF(i,j))W(i,j) \\ &+ B^{LH}.\beta.NVF(i,j)W(i,j) \\ I_l'^{HL}(i,j) &= I_l^{HL}(i,j) + A^{HL}.\alpha(1-NVF(i,j))W(i,j) \\ &+ B^{HL}.\beta.NVF(i,j)W(i,j) \end{aligned} \quad (1)$$

where $I_l'^{LH}$, $I_l'^{HL}$ are watermarked transform coefficients, A^{LH} B^{LH} , A^{HL} B^{HL} denote the watermark strengths of texture and edge regions and flat regions at I_l^{LH} and I_l^{HL} sub-bands. α , β are smoothing factors at the texture regions and flat regions, β is fixed to 0.1 where α is adapted to control the smoothness.

In this proposal the strengths of watermark A^{LH} and A^{HL} are calculated by the mean of the absolute value at I_l^{LH} and I_l^{HL} sub-bands, where $A^{LH} = \mu|I_l^{LH}|$, $A^{HL} = \mu|I_l^{HL}|$, $B^{LH} = A^{LH}/10$, and $B^{HL} = A^{HL}/10$, μ denotes the mean value. The watermark image I' is obtained by the inverse DWT.

In order to apply ICA for a blind watermark detection, the embedding process need to create a de-mix key for the detection phase. The following equation is used to create the de-mix key:

$$\begin{aligned} I_l^{*LH}(i,j) &= I_l^{LH}(i,j) + A^{LH}\alpha_k.K(i,j) \\ I_l^{*HL}(i,j) &= I_l^{HL}(i,j) + A^{HL}.\alpha_k.K(i,j) \\ demix\_key(i,j) &= I_l^{*HL}(i,j) + I_l^{*HL}(i,j) + K(i,j) \end{aligned} \quad (2)$$

where I_l^{*LH} , I_l^{*HL} are the wavelet transform coefficients at I_l^{LH}, and I_l^{HL} subbands which are embedded the random key, K, and $\alpha_k = 0.5$ is used as strength of the mixture. Using only the de-mix key and the key, the owner can claim the ownership on any watermarked and any copy versions. The original image is kept in secret and the de-mix key is used for user detection. Fig 3 shows the embedding scheme, and the de-mix key is created during the embedding phase.

3. ICA BASED WATERMARK EXTRACTION

3.1. Independent component analysis

Independent Component Analysis (ICA) is a method for extracting independent sources, given only mixtures of the unknown sources. This method has a wide range of applications in signal and image processing issues. It is noted that the watermark, and the original can be regarded as unknown sources and the watermarked is a mixture. By creating different mixtures, one can perform ICA to extract the watermark. Several algorithms for ICA have been introduced recently [4] [5] [6] [7]. This paper applied fast ICA algorithm for the watermark extraction: fixed point or fast ICA algorithm was originally developed by Aapo Hyvärinen and Erkki Oja [4] which is based on the following two stages: first is PCA whitening process of the input and the next is the fast ICA algorithm by using the fourth-order statistics of the signal. Fast ICA is chosen because it has a number of good properties compared with the existing methods for ICA, which include fast convergence, easy to implement and suitable for watermark application.

3.2. Watermark extraction

The usefulness of this new proposed method is that it does not require the knowledge of the watermark, and other parameters such as original image, strengths, smooth factors or any others. By using ICA, one can identify their ownership of any copy version of the watermarked image.

The following step is used to extract the watermark:

Step 1: The stego (watermarked) image is decomposed through DWT in l levels (l = 1, 2, 3), in order to obtain the wavelet coefficients at $I^{\#LH}_{l}$ and $I^{\#LH}_{l}$ sub-bands.

Step 2: From $I^{\#LH}_{l}$ and $I^{\#HL}_{l}$ sub bands, create mixture signals to input ICA.

$$\begin{aligned} X'_1 &= I_l^{\#LH} + demix\_key \\ X'_2 &= I_l^{\#HL} + demix\_key \\ X'_3 &= I_l^{\#LH} + I_l^{\#HL} \\ X'_4 &= demix\_key + K \end{aligned} \tag{3}$$

where X_1, X_2, X_3, X_4 are mixtures observations of the wavelet transform coefficients of the original (I_l^{LH} , I_l^{HL}), the watermark and the key. They also can be rewritten as:

$$\begin{aligned} X'_1 &= a_{11}I_l^{LH} + a_{12}I_l^{HL} + a_{13}w + a_{14}k \\ X'_2 &= a_{21}I_l^{LH} + a_{22}I_l^{HL} + a_{23}w + a_{24}k \\ X'_3 &= a_{31}I_l^{LH} + a_{32}I_l^{HL} + a_{33}w + a_{34}k \\ X'_3 &= a_{41}I_l^{LH} + a_{42}I_l^{HL} + a_{43}w + a_{44}k \end{aligned} \tag{4}$$

where $a_{i,j}(i = 1,2,3,4; j = 1,2,3,4)$ are the unknown mixture parameters. As describe earlier, it is possible to perform ICA for those mixtures to extract the watermark.

Step 3: The above four mixtures are input to fastICA [4] algorithm and the watermark W is extracted from those mixtures. The copyright can be claimed from the watermarked image or the copy versions of the watermarked image.

Fig.4 shows the ICA blind watermarked extraction method based on the de-mix key.

After extraction of the watermark, users can compare the extracted one with the reference watermark subjectively. A similarity measurement of the extracted, $W'(i,j)$, and the reference watermarks, $W(i,j)$, can be defined by the normalized correlation (NC):

$$NC = \frac{\sum_{i=1}^{M_w}\sum_{j=1}^{N_w}\left[W(i,j)W'(i,j)\right]}{\sum_{i=1}^{M_w}\sum_{j=1}^{N_w}\left[W(i,j)\right]^2} \tag{5}$$

The value of NC lies in $[0,1]$, and if we acquire higher NC values, the embedded watermark is more similar to the extracted one.

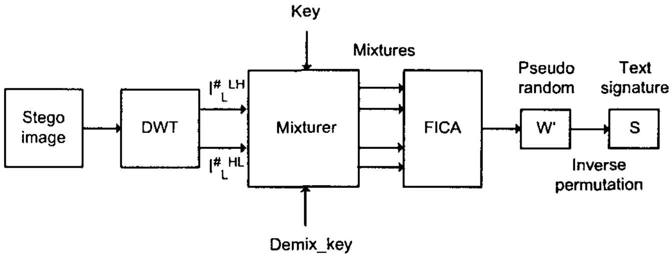
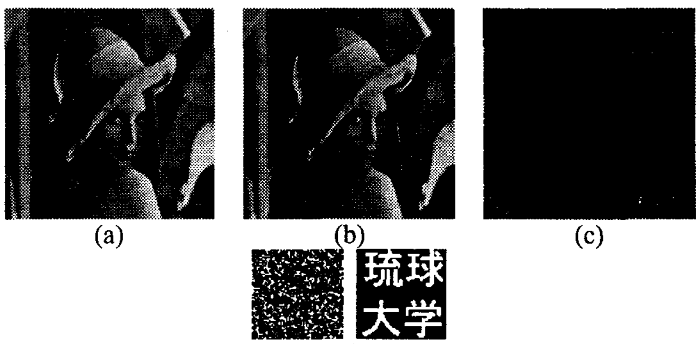


Fig.4. The proposed extraction scheme

4. EXPERIMENTAL RESULTS

The proposed watermarking method was tested on "Lena" image with size 512x512. Haar filters have been used for computing DWT with 3 levels where the highest sub-bands I_3^{LH} , I_3^{HL} are used for embedding watermark. The Japanese text signature meaning "*University of the Ryukyus*" of size 64x64 is used as watermark, and α is set to 0.2 to ensure the invisibility. Fig 5(a) and 5(b) show the original and watermarked image, it is evidently undistinguishable (PNSN = 49.30 dB). The effectiveness of NVF function on the watermarked image is shown in Fig 5(c). It can be confirmed that the watermark in texture and edge areas are stronger than in flat regions. Fig 5(d) shows extracted 64x64 random sequences and its inverse permutation.



(a) (b) (c)

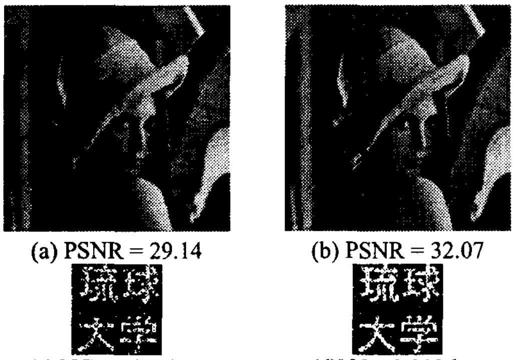
(d)

Fig.5 (a) The original "Lena" image, (b) Watermarked "Lena" with PSNR = 49.30 (c) Difference between the original "Lena" and the watermarked "Lena" (d) Extracted watermark by ICA before permutation and after permutation (NC = 1.00).

Fig.6 shows the watermarked image filtered with low-pass and median filters, its extracted watermark, PSNR, and NC values. Fig.7 shows a cropping attacked "Lena" in experiments with the corresponding PSNR, and NC of cropped results: (7a) cropped 25% lower-left quarter, (7b) cropped 45% of the surroundings and replaced with pepper image, 7(c) 7(d) the corresponding ICA watermark extracted and NC results.

Fig.8 shows the extracted results from Jpeg compression, JPEG-2000 compression [10], Wavelet-SPIHT (set-partitioning in hierarchical trees) compression [9], Embedded Zero-tree Wavelet compression (EZW) [8] and PCA based compression. The PCA based compression is applied on each

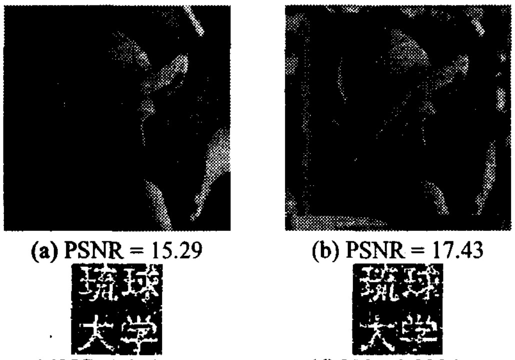
8x8 sub-block of the image. The images reconstructed by retaining only coefficients corresponding to the energy components of the sub-block.



(a) PSNR = 29.14 (b) PSNR = 32.07

(c) NC =0.8773 (d)NC =0.9136

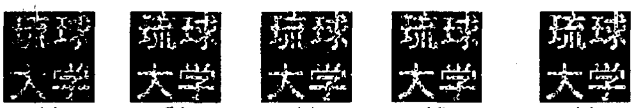
Fig.6. (a) (b) Watermarked image "Lena" low-pass and median filtered 5x5 and PSNR results, (c) (d) The corresponding ICA watermark extracted and NC results.



(a) PSNR = 15.29 (b) PSNR = 17.43

(c)NC=0.8427 (d) NC= 0.8334

Fig.7. The cropped versions "Lena" in experiment and the PSNR, NC of cropped results



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

Fig.8. (a) The extracted watermark from jpeg (quality factor = 15%), (b) jpeg 2000 (rate = 0.4), (c) SPITH (rate = 0.4), (d) EZW (rate = 0.4), and (e) PCA based compression (retain coefficients = 17) of the watermarked image "Lena".

Table 1 shows PSNR, NC values of Jpeg, Jpeg 2000, SPITH, EZW and PCA compression attacks corresponding to Fig.8. Table 2 shows PSNR and NC values after addition of random noise, denoise by Wiener filter, and resize of the watermarked version "Lena" to difference size.

Table 1. PSNR, NC values of Jpeg, Jpeg 2000, SPITH, EZW and PCA compress attacks. The higher NC values indicate , the more robustness

Attacks	JPEG	JPEG2000	EZW	SPIHT	PCA
PSNR	32.4162	36.0245	36.5377	35.7217	36.8136
NC	0.7628	0.8647	0.8335	0.8161	0.9314

Table 2. NC values vs. the attacking methods. The higher NC values indicate , the more robustness

Attacks	PSNR(dB)	NC
Adding random noise Power = 200	28.1321	0.8163
Random noise + denoising by Wiener filter	34.0117	0.8089
Resize to 448x448	37.6637	0.9702
Resize to 640x640	39.5429	0.9585
Resize to 256x256	32.3801	0.8628

5. CONCLUSION

In this paper, we proposed a new image watermark technique based on ICA to extract the watermark. By ICA, even without information on the embedding such as original image, strength, watermark as well as threshold, the watermark can perfectly extracted. Our algorithm is based on the wavelet domain that is suitable for the incoming Jpeg 2000 compression algorithm. The experimental results show the proposed embedding technique can survive under almost all compression domains such as DCT, DWT and PCA based compression. The robustness was also checked under various types of attacks including cropping of an image, low-pass and median filtering, adding noise, image resize. Further work will concentrate on enhancement the proposed algorithm to protect the original source more effectively

REFERENCES

[1] M. Barni, F. Bartolini, and A. Piva, "Improved Wavelet-Based Watermarking Through Pixel-Wise Masking," IEEE Trans., Image Process., vol.10, no. 5, pp. 783-791, 2001.

[2] S. Voloshynovskiy, A. Herrigel, N. B., and T. Pun, "A stochastic approach to content adaptive digital image watermarking," Lecture Notes in Computer Science, vol.1768, pp. 212-236, September 2000.

[3]Christine I. Podilchuk, Wenjun Zeng, "Image-Adaptive Watermarking Using Visual Models," IEEE Journal Selected Areas of Communications (JSAC), vol.16, No.4,pp. 525-539, 1998.

[4] E. Bingham and A. Hyvärinen, "A fast fixed-point algorithm for independent component analysis of complex-valued signals," Int. Journal of Neural Systems, vol.10, no.1, pp.1-8, 2000.

[5] Bell A.J. and Sejnowski T.J, "An information maximization approach to blind separation and blind deconvolution," Neural Computation, vol.7, no.6, pp.1129-1159, 1995.

[6] J.-F. Cardoso, "High-order contrasts for independent component analysis," Neural computation, Vol.11, no 1, pp.157-192, 1999.

[7] S. Cruces, L. Castedo, A. Cichocki, "Robust blind source separation algorithms using cumulants," Neurocomputing, vol.49, pp.87-118, 2002.

[8] M Shapiro, "Embedded image coding using zerotree of wavelet coefficients," IEEE Trans., Signal Proc., vol.41, no.12, 1993.

[9] Amir Said and William A. Pearlman "A New Fast and Efficient Image Codec Based on Set Partitioning in Hierarchical Trees," IEEE Trans., Circuits and Systems for Video Tech., vol. 6, pp.243-250, 1996.

[10] C. Christopoulos, A. Skodras, and T. Ebrahimi, "The JPEG2000 still image coding system: An Overview," IEEE Trans., Consumer Electronics, Vol. 46, No. 4, pp. 1103-1127, 2000.