Effective Approach for Detecting Digital Image Watermarking via Independent Component Analysis

Lisha Sun[1], Weiling Xu[1], Zhancheng Li[1], M. Shen[1], and Patch Beadle[2]

[1] Key Lab. of Guangdong, Shantou University, Guangdong 515063, China
mfshen@stu.edu.cn
[2] School of System Engineering, Portsmouth University, Portsmouth, U.K.

Abstract. A basic scheme for extracting digital image watermark is proposed using independent component analysis (ICA). The algorithm in terms of fastICA is discussed and used to separate the watermark from the mixed sources. The behavior of the proposed approach with several robustness tests of the image watermark is also carried out to demonstrate that ICA technique could provide a flexible and robust system for performing digital watermark detection and extraction. The preliminary experimental results show that the proposed watermarking method is effective and robust to some possible attacks.

1 Introduction

In the past decade, there exist many methods developed for hiding digital image watermarks in various areas such as digital images, video and other multimedia for the purposes of copyright protection**.** The success and the effectiveness of assessing the digital watermarking methods are based on both the efficiency of the algorithms used and the abilities of resisting the possible attacks. Recently, there is a rapid growth of digital image and digital image watermark since the recent growth of network multimedia systems has met a series of problems related to the protection of intellectual property rights. Digital watermark can be regarded as a procedure of a robust and imperceptible digital code, which consists of the specified information embedded in the host signals like digital images. All types of protection systems involve the use of both encryption and authentication techniques. One of these ideas for the protection of intellectual property rights is embedding digital watermarks into multimedia data [1]. The watermark is a digital code irremovably, robustly, and imperceptibly embedded in the host data and typically contains information about origin, status, and destination of the signals. The basic principles of watermarking methods use small and pseudorandom changes to the selected coefficients in the spatial or transform domain. Most of the watermark detection schemes apply some kinds of correlating detector to verify the presence of the embedded watermarking [1].

ICA technique is a signal processing algorithm to represent a finite set of random variables as the linear combinations of independent component variables [2,3]. The ICA for digital watermarking belongs to the method of removal attack [4]. In this contribution, ICA was proposed to deal with the problem of detecting the digital image watermark and testing the robustness of the proposed scheme.

G.I. Webb and X. Yu (Eds.): AI 2004, LNAI 3339, pp. 949–954, 2004.

2 The Scheme of Digital Watermarking

Firstly, the procedure of watermarking embedding is provided. The basic idea is to add a watermark signal to the host data to be watermarked so that the watermark signal is unobtrusive and secure in the signal mixture but can be recovered from the signal mixture later on. Generally, three main topics were involved for designing a watermarking system, including design of the watermark W to be added to the host signal, the embedding method which incorporates the watermark to the host signal X to obtain the watermarked signal Y, and the proper extraction algorithm to recover the watermark information from the mixing signal. The watermark should be any signal related with a message. As a matter of fact, the differences of the watermark method are more or less dependent on the signal design, embedding, and recovery. Usually the correlation techniques are employed for watermark recovery [5]. We adopt the embedding procedure for our ICA scheme

$$W = X + aK + b * M \tag{1}$$

$$W = X + aK + bM \tag{2}$$

where X is the host data, K denotes key and the star symbol represents the convolution operation. Both M and K are inserted in the spatial domain of the X while a and b stand for the small weighting coefficients. The number of observed linear mixture inputs is required to at least equal to or larger than the number of independent sources so that the identification of ICA can be performed. Mostly, at least we need three linear mixtures of three independent sources for our purpose. Two more mixed images are generated to be added to the watermarked image W by using the key image and the original image I in which both c and d denote arbitrary real numbers:

$$W_1 = W \ , \ W_2 = W + cK \ , \ W_3 = W + dI \tag{3}$$

To apply ICA algorithm, three images above can be set as three rows in one matrix for the purpose of de-watermarking.

3 Blind Extraction

By using ICA, we desire to minimize the statistical dependence of the component of the representation [3,4]. The ICA is supposed that the time courses of activation of the sources are as statistically independent as possible. Most ICA is performed using information-theoretic unsupervised learning algorithms [4,5]. In this contribution, the fixed-point algorithm is adopted for detecting digital image watermark in two stages. First of all, the procedure of principal component analysis was used for whitening such that the whitened data matrix has the following form [6]

$$\mathrm{Y} = \Lambda_s^{-1/2} U_s^T R \tag{4}$$

Where Λ_s denotes the diagonal matrix containing k eigenvalues of the estimated data correlation matrix, and Us is the matrix containing the respective eigenvectors in the same order. Thus from the rank of the diagonal matrix, the number of sources or independent components can be determined. Secondly, higher-order statistics (HOS) and their characteristics [7,8] were used for our problem. After finishing the procedure of whitening, the fastICA algorithm in terms of HOS can be summarized as the following three stages [9,10]: First, we need to choose an initial vector w(0) randomly which is normalized to be unit norm. The send stage is to estimate one ICA basis vector by using the following fixed-point iteration procedure:

$$w(k) = Y[Y^T w(k-1)]^3 - 3w(k-1) \tag{5}$$

where $(\cdot)^3$ means the element-wise operation. Finally, w(k) is normalized in terms of dividing it by its norm. When w(k) is not converged, we need to go back to the second stage. If we can project a new initial basis vector w(0) onto the subspace which is orthogonal to the subspace spanned by the previously found ICA basis vectors, and follow the same procedure, other ICA basis vectors can be estimated sequentially.

4 Experimental Results

In this section, ICA is applied with some simulations to show the validity and feasibility of the proposed scheme. Both watermark detection and extraction are investigated. Fig.1 shows an example of watermark extraction. The performance of watermark extraction is evaluated by calculating the defined normalized correlation coefficient [11,12]:

$$r = \frac{\sum_{i=1}^{L} m(i)\,\hat{m}(i)}{\sqrt{\sum_{i=1}^{L} m(i)^2 \sum_{i=1}^{L} \hat{m}(i)^2}} \tag{6}$$

where L denotes the total number of pixels of the image, and both m and $\hat{m}$ represent the original and the extracted watermark sequences with zero-mean values, respectively. The value range of r is between minus one and unity. The unit r means that the image extracted perfectly matched the original. The minus sign indicates that the extracted image is a reverse version of its original image. To evaluate the performance of the example in Fig.1, the normalized correlation coefficients between the original and the extracted images were estimated with the host image of 0.9991, the key image of about unity and the watermark of 0.9989, which proves that the fast ICA algorithm effectively separates the images from the mixture signal.

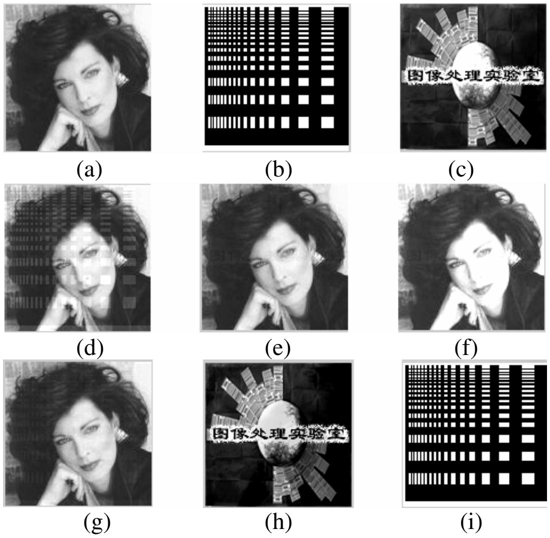

Fig. 1. (a) original Debbie image, (b) key image, (c) watermark, (d) watermarked image, (e) and (f) generated mixture images, (g) extracted Debbie image. (h) extracted watermark, (i) extracted key

5 Test of Robustness

Test of watermark attack is another important problem for assessing the performance of the proposed digital watermarking method [13]. The watermark attack is a procedure which can be used to evaluate the robustness of the presented watermarking scheme. The watermarking system should be robust against data distortions introduced through standard data processing and attacks. There are several watermark attack techniques such as simple attacks, removal attacks and detection-disabling attacks. In this section, we focus on testing the attack performances under the conditions of noise addition, the image compression and the filtering. Firstly, the test of the noise addition was investigated. The watermarked Cameraman image is corrupted by the Gaussian noise. One simulation was carried out and shown in Fig. 2. Note that the maximum acceptable noise level is limited by comparing the energy strength of the embedded watermark. When the additive noise energy level goes up to 40-50 times higher than the energy level of the text watermark, the simulation shows that the watermark become unpreventable. Next, the operation-compression is employed to test the watermarked image by using the Lenna image. The compressed format is JPEG and the compressed proportion is set with 8:1. Fig. 3 (a) and (b) show the original images. The extracted Lenna image and the watermark image in terms of the proposed algorithm are shown in Fig. 3 (c) and (d). The test results via JPEC compression demonstrate the success of the presented ICA method in extracting the watermark even after compression attacks. Finally, the attack of low pass filtering was carried out. Fig. 4 (a) and (b) give two watermarked Debbie images filtered with a 2D low-pass Gaussian and a 2D average filter of size 5x5, respectively. The text watermark was shown in Fig. 4 (c). The watermarked Debbie image filtered with a low pass average filter was demonstrated in Fig. 4 (d) while the extracted watermark image was given in Fig. 4 (e). It can be seen that the ICA scheme can well survive these types of low pass filtering attacks.

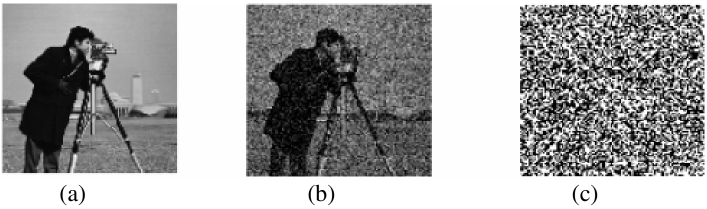

Fig. 2. Test of the strong noise attack. (a) original Cameraman image. (b) extracted Cameraman image. (c) extracted the watermark noise

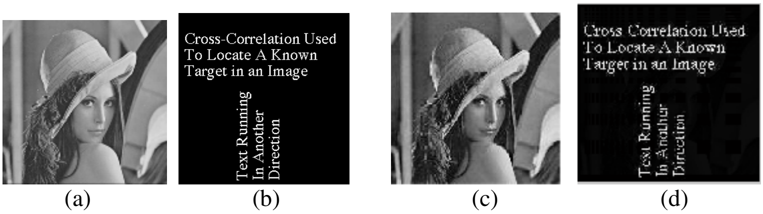

Fig. 3. Illustartion of the robustness of ICA demixing ability with respect to JPEG compression. (a-b) the original images. (c-d) the extracted image from compressed mixtures of the originals

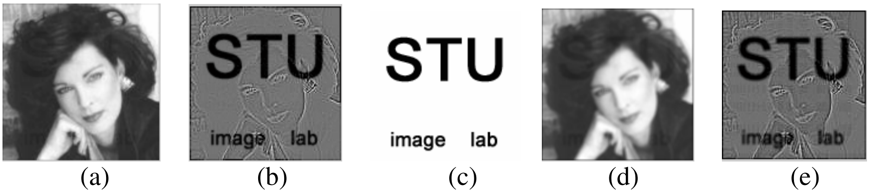

Fig. 4. The attack test with 2-D low-pass Gaussian filtering and two dimensional average filtering

6 Conclusions

We have presented a new scheme for the detection and the extraction of digital image watermarking based on independent component analysis. The fastICA algorithm was discussed and adopted to the problem of image processing. In addition, the ICA was used to investigate the robustness of the proposed procedure of digital image watermarking. Several aspects of attacks were also tested. The preliminary experimental results demonstrate the success of ICA algorithm in performing the watermark detection and extraction.

Acknowledgements

The research was supported by the Natural Science Foundation of China (60271023), the Natural Science Foundation of Guangdong (021264 and 32025), and the Key Grant of the Education Ministry of China (02110).

References

1. Hartung F. and Kutter M.: Multimedia Watermarking Techniques. Proceedings of the IEEE, Vol.87, No.7, (1999) 1079-1107
2. Comon P.: Independent Component Analysis, a New Concept? Signal Processing. Vol. 36, (1994) 287-314
3. Aapo H.: Survey on Independent Component Analysis. Neural Computing Surveys, Vol.2, (1999) 94-128
4. Cardoso J. F.: Blind Signal Separation: Statistical Principles. Proceedings of the IEEE, Vol.9, no.10, (1998) 2009-2026
5. Yu D., Sattar F., and Ma K.: Watermark Detection and Extraction Using an ICA Method. EURASIP Journal on Applied Signal Processing, (2002) 92-104
6. Petitcolas F. A. P., Anderson R. J.: Evaluation of Copyright Marking Systems. Proceedings of IEEE Multimedia Systems, Vol.1. (1999) 574-579
7. Vidal J. and et al., Causal AR Modeling Using a Linear Combination of Cumulant Slices, Signal Processing, Vol. 36. (1994) 329-340
8. Shen M, Chan F. H. Y., Sun L, and Beadle B. J.: Parametric Bispectral Estimation of EEG Signals in Different Functional States of the Brain. IEE Proceedings in Science, Measurement and Technology, Vol.147, No.6. (2000) 374-377
9. Hyvarinen A. and Oja E.: A Fast-fixed Point Algorithm for Independent Component Analysis. Neural Computation, (1997) 1483-1492
10. Hyvarinen A.: Fast and Robust Fixed-point Algorithm for Independent Component Analysis. IEEE Trans. on Neural Network, Vol. 10 (1999) 626-634
11. Kashyap R. L.: Robust Image Models and Their Applications. Advances in Electronics and Electron Physics, P. W. Hawkes, Ed., vol. 70. Academic Press (1988) 79-157
12. Juan R. and et. al.: Statistical Analysis of Watermarking Schemes for copyright Protection of Images. Proceedings of the IEEE, Vol.87, No.7. (1999) 1142-1166
13. Petitcolas F. A. P., Anderson R. J., and Kuhn M. G.: Attacks on Copyright Marking Systems. 2nd International Workshop on Information Hiding, Lecture Notes in Computer Science Vol.1525 (1998) 218-238