

Improved Watermark Extraction Exploiting Undetermined Source Separation Methods

Mohammed Khalil¹, Nawal EL Hamdouni², and Abdellah Adib¹

¹ LIM@II-FSTM, B.P. 146 Mohammedia, 20650, Morocco
{khalil,adib}@fstm.ac.ma

² LRIT -FSR, av. Ibn Battouta, B.P. 1014, Rabat, Morocco
nawaldouni@gmail.com

Abstract. In this paper, we propose an efficient blind audio watermarking detection method for spread spectrum watermarking system. The proposed method considers the detection problem of the watermark as a matter of Blind Source Separation (BSS) between the audio signal and the watermark one. Thus, we aim to improve detection step by employing Independent Subspace Analysis (ISA) and Empirical Mode Decomposition (EMD) theories. The resulting technique called UISA (Underdetermined Independent Subspace Analysis) permits extracting the embedded watermark with low binary error rate and high bit rate.

1 Introduction

Audio watermarking is a process of embedding into a host signal, a perceptually transparent digital watermark [4]. Initially, directed towards copyright protection, the problem can be seen as a data hiding transmission [12]. The audio signal is considered as a noise whereas the watermark signal is considered as data to be recovered [3]. The watermark signal carrying the information, is characterized by a low power compared to the audio signal, whose properties are very different to white gaussian noise.

Several audio watermarking methods proposed in the literature are based on spread spectrum technique [1] [3] [4] [11]. This is due to potential reliability, inaudibility, and robustness of the technique. In [3], authors present new informed embedding strategies that ensure the local inaudibility constraint and maximizes the robustness of the transmission to channel disturbances. In [4], effective embedding and extracting mechanisms based on direct spread spectrum techniques are discussed. Moreover, various efforts on spread spectrum watermarking have dealt with detecting procedures using filters at the reception. In [11], authors propose an improved technique for detecting audio watermark by introducing spectral envelope filter. The watermarking system proposed in [1], uses wiener equalizer in order to improve the detection of the watermark at the reception.

Nevertheless, the low power of the watermark signal and its non-stationary nature constitute a great issue to conventional filter approaches [12]. BSS methods would be a good solution to overcome equalization problems. Until now, an analysis about BSS performance has not been provided sufficiently in watermarking

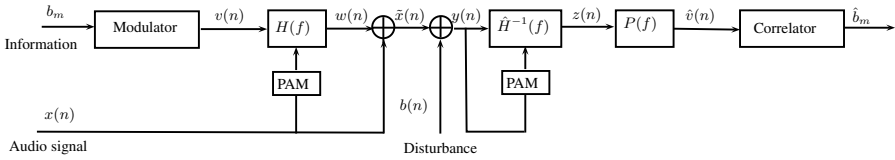


Fig. 1. Generic spread spectrum watermarking system

context [14] [13]. In this study, we restrict our attention on BSS techniques to improve the effectiveness of detecting in audio. We focus on designing a blind detector using a new BSS technique called UISA [7]. This technique comes from two existing methods used in a context different to watermarking : EMD [8] and Independent Subspace Analysis (ISA) [6]. The proposed algorithm allows to recover two sources (the host signal and the watermark signal) from only one signal, the watermarked one.

After analyzing the generic spread spectrum watermarking scheme in the next section, we introduce the new watermark extraction based on UISA technique in section III. Experimental and various test results discussed in section IV, allow us to evaluate the performances of the new approach. Finally, the conclusion ends this paper.

2 Generic Spread Spectrum Watermarking Scheme

In this section, we will present a brief overview of the generic spread spectrum watermarking scheme depicted in Fig. 1. The watermarking system is seen as a very particular communication chain, where the noisy channel is considered as the host signal. The emitter has to insert the watermark in the audio signal under the inaudibility constraint. At the reception, the hidden information will be extracted from only the watermarked audio signal.

2.1 Embedding Process

A spread spectrum watermark signal $v(n)$ is obtained by modulating a watermark sequence of M bits $b_m \in \{0, 1\}$ with a spreading signal $u(n)$ composed of N_b samples ($u(0), u(1), \dots, u(N_b - 1)$)

$$v(n) = \sum_{m=0}^{M-1} a_m u(n - mN_b) \tag{1}$$

where a_m is a symbol in $\{-1, 1\}$ given by $a_m = 2b_m - 1$.

The modulated signal $v(n)$ can't be embedded directly into the audio signal. Indeed, the power of the modulated signal σ_v^2 is very small compared to the host signal power. For this, a shaping filter $H(f)$ is designed by a Psychoacoustic Auditory Model (PAM) in order to increase the Signal-to-noise ratio (SNR) of the watermark, and make it as inaudible as possible [12].

2.2 Equalization Techniques

In the generic scheme depicted in Fig. 1, the watermark detection step is considered as a blind equalization problem which aims to improve the detection performance in terms of bit error rate. In the first time, we introduce the zero-forcing equalizer, then, we present the prediction filter. Note that the generic scheme that includes predicting filter, after the zero forcing equalizer, will be considered as the improved generic system. The generic system and the improved one will be compared to the proposed approach (presented in the next section) in terms of detection reliability.

Zero-Forcing Equalizer. The zero-forcing equalizer cancels the inter-symbol interference by filtering the watermarked signal with the inverse of the channel [3]. However, since the audio signal $x(n)$ used to derive $H(f)$ is not available at the receiver, an estimate $\frac{1}{\hat{H}(f)}$ is computed by psychoacoustic modeling of the watermarked signal $y(n)$.

Prediction Filter. The predictor filter $P(z)$ is equivalent to whitening filter of the signal $z(n)$ [12] and it is written as :

$$P(z) = \frac{1}{\beta_0}(1 + \alpha_p(1)z^{-1} + \dots + \alpha_p(M)z^{-M}) \quad (2)$$

where $\beta_0 = \sqrt{\sigma_e^2}$ and the power of the prediction error is given by :

$$\sigma_e^2 = \sum_{i=0}^M \alpha_p(i)r_z(i). \quad (3)$$

and $r_z = [r_z(1), \dots, r_z(M)]^t$ is the autocorrelation vector and α_p are the optimal coefficients.

3 The Proposed Watermark Extraction Method

Hereafter, we try to substitute, in the detection step, the equalizer procedures by the BSS ones. Firstly, we have adapted two BSS methods to the audio watermarking context : blind audio sources separation using modal decomposition (BASSMD) [5] and independent subspace analysis (ISA) technique [6]. The performances of these two techniques will be compared with those of the proposed one.

Indeed, classical BSS techniques aim at estimating N original source signals from P observations of their mixture. In our case we have two sources ($x(n)$ and $w(n)$) and only one observation : the watermarked one $y(n)$. Recovering the sources with less sensors than mixtures seems a much harder problem especially when very few knowledge about the sources and the mixture process are available. To overcome this difficulty, the idea is to use the EMD decomposition to exploit a large number of IMFs which leads us to transit from an under-determined case to an above-determined¹ case.

¹ There is now more sensors rather than one represented by the watermarked signal.

As audio signals are mostly non-stationary, then EMD is used here as a suitable tool for its decomposition. Moreover, EMD uses only a single mixture to extract IMFs components. So, the resulting mixed signal $y(n)$ is passed through EMD technique to attain the resulting IMFs. This allows us to switch from an observation $y(n)$ to a set of observations which are IMFs of $y(n)$. Then, the resulting temporal matrix of IMFs is used to derive a set of independent basis vectors by applying a ISA technique [6] [7]. The proposed UISA extraction proceeds as follows :

1. Initially, the watermarked signal $y(n)$ is converted to a time-imfs representation by applying **EMD** decomposition of $y(n)$, which build the observation matrix **IMF** :

$$\mathbf{IMF} = [IMF_1(n), IMF_2(n), \dots, IMF_M(n)] \tag{4}$$

where M is the number of *IMFs*. The dimensionality of $IMF_m(n)$ is $1 \times L$ where L is the length of data.

2. Apply PCA algorithm by employing SVD [6] [7] according to the following equation:

$$\mathbf{IMF} = \mathbf{U}\Sigma\mathbf{V}^T \tag{5}$$

where \mathbf{U} and \mathbf{V} are the orthogonal matrices with orthogonal columns with respectively $L \times L$ and $M \times M$, and Σ is the matrix which contains M singular values.

3. Use a reduced set of ρ basis vectors selected from \mathbf{U} , i.e $\mathbf{U}_{L \times \rho} = [\mathbf{u}_{L \times 1}, \mathbf{u}_{L \times 2}, \dots, \mathbf{u}_{L \times \rho}]$ by using the first ρ singular values defined by the information contained in the nonzero singular values [9].
4. Apply ICA to express the observation vector (set of bases derived by PCA) $\mathbf{U}_{L \times \rho}$ as the product of mixing matrix \mathbf{Q} and statistically independent vector \mathbf{s} :

$$\mathbf{U}_{L \times \rho} = \mathbf{Q}\mathbf{s} \tag{6}$$

where \mathbf{Q} is a $\rho \times P$ (pseudo-) invertible matrix with orthogonal columns, \mathbf{s} is a random vector of P source components.

5. Build the independent subspaces using the equations below:

$$\mathbf{Y} = \mathbf{A}\mathbf{u} = \mathbf{A}\mathbf{Q}\mathbf{s} \tag{7}$$

$$\mathbf{Z} = \mathbf{v}\Sigma\mathbf{A} \tag{8}$$

6. Construct an imf matrix of an independent subspace (sources) obtained from:

$$\mathbf{IMF}^j = y_j z_j^T \tag{9}$$

where $y_j \in \mathbf{Y}$ and $z_j \in \mathbf{Z}$.

7. Implement a reversible process of EMD decomposition, so that the time domain signal can be calculated from the simple summation of *IMFs*, for each *IMFs* matrix, IMF_j , of an independent subspace.
8. Clustering of components to sources resulting from the k-means algorithm.

4 Performance Evaluation of the Proposed Method

In this section, we present various test results in order to point out the assets of the proposed extraction technique. The performances of this method will be compared, firstly with the generic scheme and thereafter with two existing BSS methods : blind audio sources separation using modal decomposition (BASSMD) independent subspace analysis (ISA).

4.1 Experimental Protocol

The most important requirement for audio watermarking schemes is the inaudibility of the embedded watermark. Then, the detection performance are evaluated in terms of bit error rate (BER) resulting from 30 Monte-Carlo simulation runs, performed on detection of random binary sequences. The error probability detection can be estimated by the BER with a precision $p = \sqrt{BER(1 - BER)/B}$ with B is the full number of emitted bits. It is noticed that the performances of the system depend also on the host audio signal types, so we selected a set of 5 digital audio signals sampled at $F_s = 44.1kHz$ of various styles (instrumental music, sung voice and speech). The watermarks are modulated at 1 bit/symbol.

Hereafter, we consider some disturbances among the more adapted to our context, namely **MPEG** compression, performed by an **MPEG 1 Layer III** for audio signals, at bit rates of 96 kbits/s and 64 kbits/s, an operation of linear quantization with 12 bits, the effect of white additive noise with $SNR = 30dB$, a high-pass filtering operation with cutoff frequency $8Khz$.

4.2 Test Results

All test results presented below represent the mean value over the set of the audio signals. Listening tests evaluate the transparency of the watermark that are performed on a set of three listeners. They confirm for the watermarking system that the watermark is imperceptible due to the use of the spread spectrum modulation at the emitter [3].

After evaluating the inaudibility of the embedded watermark into the audio signal, it is necessary to measure the performances of the new method in terms of detection reliability. Fig. 2 shows the gain provided by the proposed UISA technique, compared to the generic detection using only zero-forcing equalizer. We can also notice that the performance of the two underterminated source separation (BASSMD and ISA) are not much improved compared to using only the generic detection.

The insufficiencies of the generic system in terms of BER make us to achieve a cascade realization of zero forcing equalizer and predicting filter (improved generic system) [12]. Fig. 3 shows that the results are very similar between improved generic system and UISA technique. In order to decrease more the BER we combine the prediction filter with the UISA technique and we remark then that the detection reliability is improved.

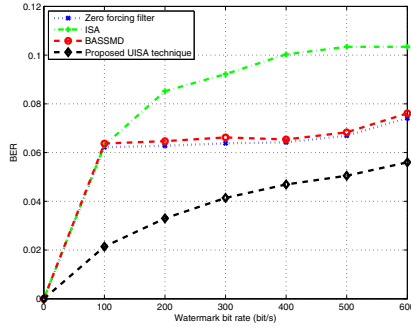


Fig. 2. Detection performance comparison between generic system and BSS techniques in the case of zero-forcing equalizer

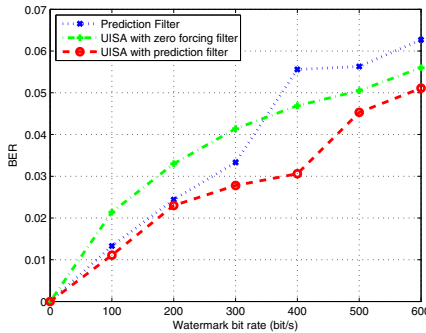


Fig. 3. Detection performance comparison between improved generic system and UIISA techniques in two cases : zero-forcing equalizer and adding prediction filter

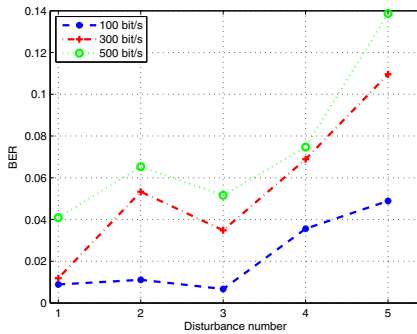


Fig. 4. Robustness of the UIISA technique with respect to channel disturbances. Disturbances are numbered as follows: 1: high-pass filtering, 2: quantization with 12 bits, 3: white noise adding with SNR=30 dB, 4 : MPEG 1 Layer III 64 kbits/s, 5 : MPEG 1 Layer III 64 kbits/s.

Watermarking system robustness against disturbances is exhibited in Fig. 4. The various non-desynchronizing disturbances give the obtained BER when the watermark bit rate is fixed at 100 bit/s, 300 bit/s and 500 bit/s. The results are satisfying and the system prove its robustness against various disturbances.

In order to measure the distortion between the original watermark signal and the estimated one, we use the normalized mean squares estimation errors (NMSE) [5] and the improvement of signal-to-noise ratio (ISNR) [15] as the quantitative measure of separation performance. The results of applying the algorithms to the watermarked signal (the mixture) are shown in Table 1 in terms of NMSE and ISNR.

Table 1. Experimental separation results (in terms of NMSE and ISNR of the watermark signal) of the proposed sources separation algorithms

Performances	ISA	BASSMD	UISA
NMSE (dB)	-3.0645	-12.1627	-19.1373
ISNR (dB)	4.3886	6.4508	12.4839

5 Conclusion

This paper presents a new approach of BSS techniques in spread spectrum watermarking scheme viewed as a hidden data transmission system. The great problem is to improve detection reliability in terms of BER. The proposed UISA method overcomes the generic system and the two BSS techniques considered in this paper. The new method improves considerably the data transmission reliability in the audio channel at lower as at higher bit rates.

References

1. Larbi, S., Jaidane, M., Moreau, N.: A new Wiener filtering based detection scheme for time domain perceptual audio. In: IEEE Int. Conf. on Acoustics, Speech and Signal Processing, Montréal (2004)
2. El Hamdouni, N., Adib, A., Larbi, S., Turki, M.: Hybrid embedding strategy for a blind audio watermarking system using EMD and ISA technique. In: ISCCSP 2010, Limassol, Cyprus (2010)
3. Baras, C., Moreau, N., Dymarski, P.: Controlling the Inaudibility and Maximizing the Robustness in an Audio Annotation Watermarking System. IEEE Transactions on Audio, Speech, and Language Processing 14(5) (September 2006)
4. Kirovski, D., Malvar, H.S.: Spread-Spectrum Watermarking of Audio Signals. IEEE Transactions on Signal Processing: Special Issue on Data Hiding, December 2 (2002)
5. Aïssa-El-Bey, A., Abed-Meraim, K., Grenier, Y.: Underdetermined blind audio source separation using modal decomposition. EURASIP Journal on Audio, Speech & Music Processing (2007)

6. Casey, M.A., Westner, A.: Separation of Mixed Audio Sources By Independent Subspace Analysis. In: ICMC, Berlin, Germany (2000)
7. El Hamdouni, N., Adib, A.: Single Mixture Audio Sources Separation Using ISA Technique in EMD Domain. In: ISIVC 2010, Rabat, Maroc (2010)
8. Huang, N.E., Shen, Z., Long, S.R., Wu, M.L., Shih, H.H., Zheng, Q., Yen, N.C., Tung, C.C., Liu, H.H.: The empirical mode decomposition and hilbert spectrum for nonlinear and non-stationary time series analysis. *R. Soc. London A* 454, 903–995 (1998)
9. Uhle, C., Dittmar, C., Sporer, T.: Extraction of drum tracks from polyphonic music using independent subspace analysis. In: *Int. Symp. Independent Compon. Anal. Blind Signal Separation, ICA* (2003)
10. Frank, I.E., Todeschini, R.: *The data analysis handbook*. Elsevier Sci. Pub. Co. (1994)
11. Jung, S., Seok, J., Hong, J.: An Improved Detection Technique for Spread Spectrum Audio Watermarking with a Spectral Envelope Filter. *ETRI Journal* 25(1) (February 2003)
12. Larbi, S.: *Structures d'égalisation en tatouage audionumérique*, Thèse présentée pour obtenir le grade de docteur de l'Ecole Nationale d'Ingénieurs de Tunis et de l'Ecole Nationale Supérieure des Télécommunications de Paris, Mars (2005)
13. Parvaix, M., Girin, L., Brossier, J.-M.: A Watermarking-based method for single channel audio source separation. In: *ICASSP, Taiwan, China* (2009)
14. Malik, H., Khokhar, A., Ansari, R.: Improved watermark detection for spread-spectrum based watermarking using independent component analysis. In: *DRM 2005 Proceedings of the 5th ACM Workshop on Digital Rights Management*. ACM, New York (2005)
15. Vincent, E., Gribonval, R., Fevotte, C.: Performance measurement in blind audio source separation. *IEEE Trans. Speech Audio Process* (2005)