



# Audio Watermarking Using Independent Component Analysis

Jongwon Seok, *Member, KIICE*

Department of Information and Communication, Changwon National University, Changwon 641-773, Korea

## Abstract

This paper presents a blind watermark detection scheme for an additive watermark embedding model. The proposed estimation-correlation-based watermark detector first estimates the embedded watermark by exploiting non-Gaussian of the real-world audio signal and the mutual independence between the host-signal and the embedded watermark and then a correlation-based detector is used to determine the presence or the absence of the watermark. For watermark estimation, blind source separation (BSS) based on independent component analysis (ICA) is used. Low watermark-to-signal ratio (WSR) is one of the limitations of blind detection with the additive embedding model. The proposed detector uses two-stage processing to improve the WSR at the blind detector; the first stage removes the audio spectrum from the watermarked audio signal using linear predictive (LP) filtering and the second stage uses the resulting residue from the LP filtering stage to estimate the embedded watermark using BSS based on ICA. Simulation results show that the proposed detector performs significantly better than existing estimation-correlation-based detection schemes.

**Index Terms:** Audio, Independent component analysis, Linear predictive coding, Watermark

## I. INTRODUCTION

Spread spectrum (SS)-based watermarking is one of the most representative method of the blind additive embedding (AE) model, which relies on the theory of spread spectrum communication for information embedding and detection. More specifically, in case of SS-based watermarking, the host signal acts as interference at the blind detector (the host signal,  $x$ , is not used during the watermark detection process) and because the host signal has much higher energy than the watermark, host interference causes the detection performance to deteriorate at the blind detector (or detectors, hereon unless otherwise stated). Superior detection performance is one of the desirable features of the blind AE

model.

The main motivation of this paper is to design a blind detector for SS-based watermarking. Existing detectors for SS-based watermarking schemes are bounded by the host signal interference at the detector. The proposed detector intends to reduce the host signal interference by developing an estimation-correlation-based detection framework. The proposed detector, therefore, consists of two stages: 1) watermark estimation stage, and 2) watermark detection stage. The objective of the watermark estimation stage is to estimate the embedded watermark, which has a higher watermark-to-signal-ratio (WSR) than the watermarked audio. To accomplish this goal, blind source separation based on the underdetermined independent component

Received 09 April 2012, Revised 01 May 2012, Accepted 09 May 2012

\*Corresponding Author E-mail: [jwseok@changwon.ac.kr](mailto:jwseok@changwon.ac.kr)

Open Access <http://dx.doi.org/10.6109/jicce.2012.10.2.175>

print ISSN:2234-8255 online ISSN:2234-8883

© This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Copyright © The Korea Institute of Information and Communication Engineering

analysis (UICA) framework (i.e., ICA for more sources than sensors) is used for watermark estimation. To this end, we model the problem of blind watermark detection for AE as that of blind source separation (BSS) for underdetermined mixtures. To ensure better WSR at the watermark estimation stage, the watermarked audio is pre-processed to remove correlation in audio signal using linear predictive (LP) filtering. It has been shown that the received watermarked signal is an underdetermined linear mixture of the underlying independent sources obeying non-Gaussian distributions, therefore BSS based on the UICA framework can be used for watermark estimation [1].

A similarity measure based on correlation is then used to detect the presence or the absence of the embedded watermark in the estimated watermark. Performance of the proposed watermark detection scheme is evaluated using a sound quality assessment material (SQAM) downloaded from [2]. Simulation results for the SQAM dataset show that the proposed scheme performs significantly better than existing estimation-correlation-based detection schemes [2] based on median filtering and Wiener filtering.

## II. MOTIVATION

The majority of existing SS-based watermarking schemes [3] use an AE model to insert the watermark into the host audio. Mathematically, the SS-based watermark embedding process can be expressed as

$$y(n) = x(n) + w(n), n = 1, 2, \dots, N \tag{1}$$

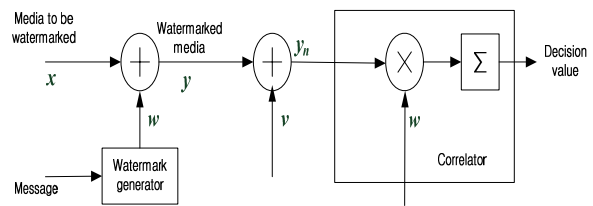
where  $y(n)$  is the watermarked audio signal in the marking space,  $x(n)$  the original signal in the marking space, and  $w(n)$  the watermark signal. It is reasonable to assume that  $x(n)$  and  $w(n)$  are zero-mean and independent and identically distributed (i.i.d.) random variables with variance,  $\sigma_x^2$ , and  $\sigma_w^2$ , respectively. It is assumed further that  $x$  and  $w$  are mutually independent. In the data hiding literature, the embedding model given by Eq. (1) is referred to as blind AE, as the embedder ignores the host signal information during the watermark embedding process.

Adversary attacks or distortion due to signal manipulations,  $v$ , can be modeled as an additive channel distortion. Therefore, the watermarked audio signal subjected to adversary attacks or channel distortions,

$$y_n(n) = y(n) + v(n), \tag{2}$$

is processed at the detector to detect the presence or the absence of the embedded watermark. The basic additive embedding and correlation based detection framework is

shown in Fig. 1.



**Fig. 1.** Semantic diagram of basic additive embedding and correlation-based detection framework.

A correlation-based detector is commonly used to detect the presence or the absence of the embedded watermark. The decision threshold,  $d$ , is obtained by correlating the received watermarked audio and the watermark sequence at the detector is given as

$$d = \frac{1}{N} \sum_{i=1}^N x(i)w(i) + \frac{1}{N} \sum_{i=1}^N w^2(i) + \frac{1}{N} \sum_{i=1}^N w(i)v(i) = d_x + d_w + d_v \tag{3}$$

where  $d_w (= \sigma_w^2)$  is the energy of the watermark,  $E_x\{d_x\} = 0$ , and  $E_v\{d_v\} = 0$ , where  $E_x\{\cdot\}$  denotes the expectation over the random variable  $x$ .

It is important to mention that detection performance of the correlation-based detector depends on the decision threshold used. Let us assume that watermark detection threshold  $T = d_x / 2$ . In this case, the probability of a false negative  $P_{fn}$  (the watermark is present, but the detector decides otherwise) equals the probability of a false positive  $P_{fp}$  (the watermark is not present but the detector decides otherwise), that is,

$$P_{fn} = P_{fp} = \frac{1}{2} \operatorname{erfc} \left( \frac{d_w}{\sqrt{8} \sigma_x} \right) \tag{4}$$

where  $\operatorname{erfc}(\cdot)$  is the complementary error function.

It can be summarized that the detection performance of a blind detector for additive watermarking schemes is inherently bounded by the host-signal interference at the detector. The motivation behind this paper is to design a watermark detector for AE with improved watermark detection performance. Towards this end, the proposed detector uses the theory of ICA by posing watermark estimation as a BSS problem from an underdetermined mixture of independent sources. The fundamentals of the ICA theory are briefly outlined in the following section followed by the details of the proposed ICA-based detector.

## A. Independent Component Analysis

ICA is a statistical framework for estimating underlying hidden factors or components of multivariate statistical data. In the ICA model, the data variables are assumed to be linear or nonlinear mixtures of some unknown latent variables, and the mixing system is also unknown [4]. Moreover, these hidden variables are assumed to be non-Gaussian and mutually independent. The linear-statistical, static ICA generative model considered in this paper, which is given as

$$X^{(i)} = AS^{(i)} + V^{(i)} : i = 1, \dots, N \quad (5)$$

where  $X^{(i)}$  represents  $N$ -realizations of an  $m$ -dimensional random vector known as the observation vector,  $S^{(i)}$  are the realizations of the  $n_1$ -dimensional source vector or the hidden variables,  $A \in R^{m \times n_1}$  is an unknown matrix also referred to as the mixing matrix, and  $V^{(i)}$  represents realizations of the noise, independent of the underlying sources,  $S^{(i)}$ .

The mixtures in which the number of observations (dimensionality of observation vector,  $X^{(i)}$ ),  $m$ , is smaller than the number of sources (dimensionality of source vector  $S^{(i)}$ ),  $n_1$ , i.e.,  $m < n_1$  is referred to as underdetermined mixtures. In this paper, we will consider such mixtures to model the blind watermark estimation problem using BSS based on the ICA framework. The BSS framework using the ICA model tries to find a linear representation of the underlying sources,  $S^{(i)} = 1, \dots, n_1$ , which are statistically independent based on the observation only. In other words, the BSS framework is intended to estimate the mixing matrix or the hidden sources  $S^{(i)}$  or both from the observation alone. As pointed out in many works such as in [1, 4], the underdetermined mixtures cannot be linearly inverted due to being bound on the rank of the mixing matrix, which makes it even more difficult to estimate the underlying sources even if the mixing matrix is known. Therefore, the problem of extracting sources from the observation following the underdetermined mixture model is nontrivial [4]. Recently, many researchers have proposed methods to solve BSS for the underdetermined mixtures problem [1, 4, 5]. For example, De Lathauwer et al. [5] have proposed a BSS scheme based on multilinear analysis, and Hojen-Sorensen et al. [1] have proposed a statistical framework based on mean-field approaches to solve BSS for underdetermined mixtures.

Before BSS based on underdetermined ICA can be used to estimate the watermark from the watermarked audio, we need to verify the following: 1) the watermarked audio is an underdetermined mixture of independent sources, and 2) the underlying sources obey a non-Gaussian distribution. It can be observed from Eq. (1) that the AE model fits into an

underdetermined linear mixture model; therefore, BSS for underdetermined mixtures can be used to estimate the embedded watermark given that the underlying latent sources ( $x$  and  $w$ ) satisfy non-Gaussianity and independence constraints. This is a realistic assumption, as multimedia data can be modeled using the non-Gaussian distribution [6]. Therefore, if an embedded watermark obeys a non-Gaussian distribution then BSS based on UICA can be used for watermark estimation from the watermarked signal [6].

## III. WATERMARK DETECTION

This section provides an overview of the proposed blind watermark detection scheme from the received watermarked audio signal obtained by additive embedding. The proposed watermark detection scheme consists of two stages: 1) the watermark estimation stage, and 2) the watermark detection stage. It was mentioned earlier that the watermark estimation stage is further divided into two sub-stages: the spectral removal stage and the source separation stage.

### A. Watermark Estimation

The goal of the watermark embedder is that the embedded watermark should survive intentional and unintentional attacks, whereas the goal of the watermark detector is to detect the embedded watermark with very low false rates in the presence of an active adversary and signal manipulations. In case of the AE model, low false rates are difficult to achieve due to strong host interference. For detector performance analysis, existing correlation-based schemes model the audio signal as a white Gaussian channel. Recent results in audio processing and compression community, however, show that samples of the real audio signals are highly correlated, which can be exploited to improve the detection performance by de-correlating the input audio before detection. The proposed detection scheme achieves this goal by applying whitening or de-correlation before watermark estimation. Simulation results presented in this paper show that the whitening before watermark estimation using ICA improves detection performance significantly. This improvement can be attributed to the fact that whitening actually increases the watermark to interference ratio and hence yields superior detection performance.

To remove the correlation in the audio signal, an autoregressive modeling named linear predictive coding (LPC) [7] can be used. The LPC method approximates the original audio signal,  $x(n)$ , as a linear combination of the past  $p$  audio samples, such that

$$x(n) \approx a_1 x(n-1) + a_2 x(n-2) + \dots + a_p x(n-p) \quad (6)$$

where the coefficients  $a_1, a_2, \dots, a_p$  are assumed to be constant over the selected audio segment. Rewriting Eq. (6) by including an error term  $e(n)$ , that is,

$$x(n) = \sum_{i=1}^p a_i x(n-i) + e(n) \quad (7)$$

where  $e(n)$  is an excitation or residual signal of  $x(n)$ . Now using Eq. (7), Eq. (1) can be expressed as

$$y(n) = \sum_{i=1}^p a_i x(n-i) + e(n) + w(n) \quad (8)$$

Likewise, watermark audio can also be expressed as,

$$y(n) = \sum_{k=1}^p a_k y(n-k) + \tilde{w}(n) \quad (9)$$

where  $\tilde{w}(n)$  is the residual signal of the watermarked audio signal. We assume that, by the characteristics of linear predictive analysis,  $\tilde{w}(n)$  has the characteristics of both  $e(n)$  and  $w(n)$ . We can consider the linear combination of the past audio sample as the estimate  $\tilde{y}(n)$ , defined as

$$\tilde{y}(n) = \sum_{k=1}^p a_k y(n-k) \quad (10)$$

Here prediction error,  $e(n)$ , can be expressed as

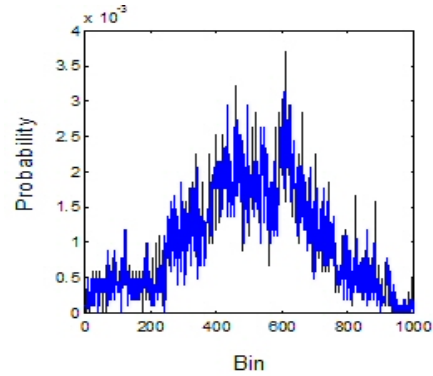
$$\begin{aligned} e(n) &= y(n) - \tilde{y}(n) \\ &= y(n) - \sum_{k=1}^p a_k y(n-k) = \tilde{w}(n) \end{aligned} \quad (11)$$

It can be observed from Eq. (11) that the estimate  $\tilde{w}(n)$  with the audio spectrum removed has the characteristics of both the excitation signal of the original audio  $x(n)$  and the watermark signal  $w(n)$ .

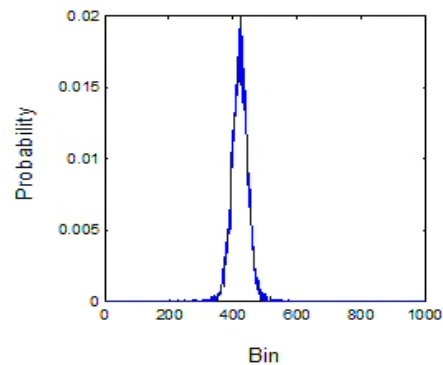
This method transforms the non-white watermarked audio signal to a whitened signal by removing the audio spectrum. It can be observed from Fig. 2 that is the empirical probability density function (pdf) of a small segment of the watermarked audio signal before LP filtering and the residual or error signal of the watermarked audio signal after LP filtering. The empirical pdf of the watermarked audio signal is clearly not smooth and has large variations due to the voiced part. On the other hand, the empirical pdf of the residual signal has a smoother distribution and a smaller variance than the watermarked audio signal.

It is important to mention that the LPC stage also improves WSR which ultimately improves the source

separation performance of the BSS used for watermark estimation. This is because, the watermark sequence is i.i.d., so the de-correlation stage does not reduce its energy in the residual signal, whereas, de-correlation does reduce audio signal energy.



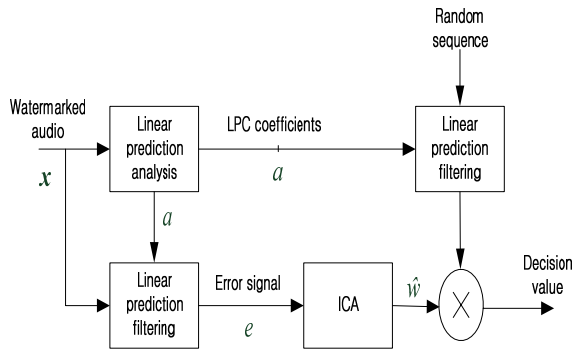
(a) watermarked audio signal



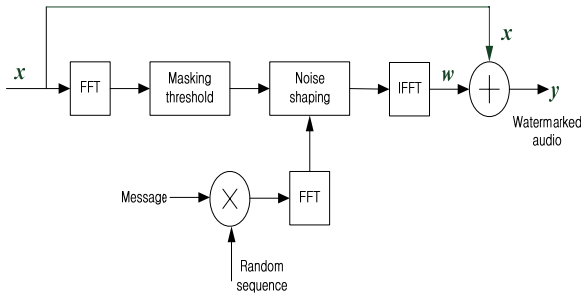
(b) Residual of watermarked audio signal

**Fig. 2.** Empirical probability density functions before and after linear prediction filtering.

The residual signal is then used to estimate the hidden watermark using BSS based on UICA. For watermark estimation, the probabilistic ICA method based on mean-field approaches is used. Superior source separation performance is the only motivation behind using of the probabilistic ICA presented in [1]. This is, however, not the limitation of the proposed scheme, as any of the BSS schemes based on UICA can be used for watermark estimation from the residual signal. Estimated sources are then correlated with the watermark,  $w$ , to determine the presence or the absence of the embedded watermark. Fig. 3 shows the block diagram of the proposed audio watermark detection scheme.



**Fig. 3.** Block diagram of the proposed watermark detection procedure. LPC: linear predictive coding, ICA: independent component analysis.



**Fig. 4.** Watermark embedding procedure. FFT: fast Fourier transform, IFFT: inverse FFT.

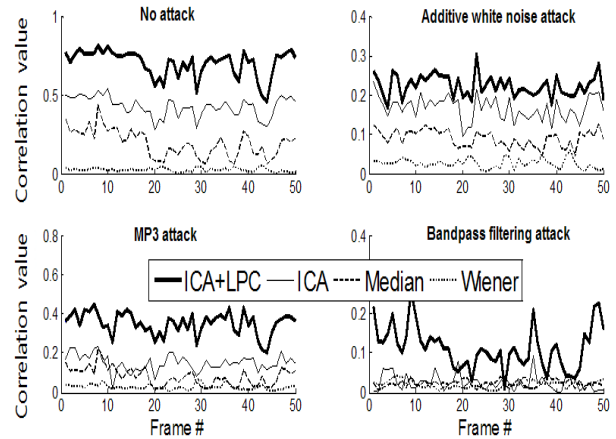
#### IV. EXPERIMENTAL RESULTS

The binary message to be embedded is first modulated by a key-dependent random sequence. The watermark is then spectrally shaped in the frequency domain according to a masking threshold estimated based on the human auditory system (HAS) ISO/MPEG-1 Audio Layer III model [8]. The motivation here is to design the weighting function that maximizes the energy of the embedded watermark subject to a required acceptable distortion. The resulting watermark is then added into the original audio signal in the frequency domain, which is then transformed to the time domain to obtain the watermarked audio. A semantic diagram of the audio watermark embedding scheme discussed above is shown in Fig. 4.

The simulation results presented in this section are based on the following system settings: 1) 44 kHz sampled and 16-bit resolution audio signals are used as the host audio, 2) a 1,024-point watermark is then embedded into four consecutive non-overlapping frames, 3) the watermarked signal is first segmented into non-overlapping frames of 4,096 samples each, then each frame is further segmented into four non-overlapping sub-frames which are then applied to the ICA block to estimate the embedded

watermark after LPC filtering. For performance evaluation, SQAM downloaded from [2] was used.

The robustness performance of the proposed the proposed watermark estimation scheme was evaluated for the following attack scenarios: 1) no adversary attack, 2) additive white Gaussian noise, 3) MP3 compression (128 kbps), and 4) bandpass filtering (2nd-order Butterworth filter with cutoff frequencies 100 and 6,000 Hz).



**Fig. 5.** Robustness performance: no Attack (top-left), additive white Gaussian noise attack (5% noise power, top-right), MP3 compression attack (128 kbps, bottom-left), and bandpass filtering attack (bottom-right). ICA: independent component analysis, LPC: linear predictive coding.

**Table 1.** Correlation values depending on the detection method using the sound quality assessment material [2] database for magnetic properties

Audio file	ICA	ICA+LPC	Median	Wiener
Bass47_1	0.449	0.775	0.315	0.046
Gspi35_2	0.519	0.787	0.206	0.056
Harp40_1	0.294	0.582	0.288	0.057
Horn23_1	0.562	0.781	0.329	0.033
Quar48_1	0.411	0.729	0.347	0.054
Sopr44_1	0.499	0.771	0.375	0.055
Trpt21_2	0.480	0.782	0.466	0.062
Vioo10_2	0.493	0.751	0.560	0.067

ICA: independent component analysis, LPC: linear predictive coding.

Detection performance of the proposed estimation-correlation based detector scheme and the existing schemes for these attacks is given in Fig 5. It is observed from Fig. 5 that, for four attack scenarios, the proposed detector outperforms the exiting detectors. In addition, detection performance of the watermark detectors under consideration for the SQAM database is given in Table 1. It can be observed from both Fig. 5 and Table 1 that the proposed

detector performs significantly better than its counterparts. Improved detection performance of the proposed detector can be attributed to its better host signal interference cancelation capability.

## V. CONCLUSIONS

In this paper, we described a new framework for estimation-correlation based detection for additive embedding. The proposed blind detection method extracts the embedded watermark signal suppressing the host signal interference at the detector. The proposed framework exploits mutual independence and non-Gaussianity of the audio signal and the embedded watermark to estimate the embedded watermark using BSS-based UICA. Experimental results showed that the proposed detection scheme is robust.

## REFERENCES

- [1] P. Hojen-Sorensen, O. Winther, and L. K. Hansen, "Mean-field approaches to independent component analysis," *Neural Computation*, vol. 14, no. 4, pp. 889-918, 2002.
- [2] SQAM, <http://tech.ebu.ch/publications/sqamcd>.
- [3] I. J. Cox, M. L. Miller, and J. A. Bloom, *Digital Watermarking*, San Diego, CA: Academic Press, 2001.
- [4] A. Hyvarinen, J. Karhunen, and E. Oja, *Independent Component Analysis*, New York, NY: John Wiley & Sons, 2001.
- [5] L. De Lathauwer, B. De Moor, and J. Vandewalle, "An algebraic ICA algorithm for 3 sources and 2 sensors," *Proceedings of the European Signal Processing Conference*, Tampere, Finland, pp. 461-465, 2000.
- [6] H. Malik, A. Khokhar, and R. Ansari, "Improved watermark detector for spread-spectrum based watermarking using independent component analysis," *Proceedings of the 5th ACM Workshop on Digital Rights Management*, Alexandria, VA, pp. 102-111, 2005.
- [7] B. S. Atal and S. L. Hanaver, "Speech analysis and synthesis by linear prediction of the speech wave," *Journal of the Acoustical Society of America*, vol. 50, no. 2, pp. 637-655, 1971.
- [8] Information Technology – Coding of Moving Pictures and Associated Audio for Digital Storage up to about 1.5 Mbits/s, ISO/IEC 11172-3:1993, 1993.



### Jongwon Seok

received his B.S., M.S. and Ph.D. degrees in electronics from Kyungpook National University, Taegu, Korea in 1993, 1995, and 1999. From 1999 to 2004, he was with the Electronics and Telecommunications Research Institute (ETRI), Korea as a senior member of the research staff. In 2004, he joined the faculty of the Department of Information and Communication at Changwon National University, Changwon, Korea, where he is currently an associate professor. His current research interests include underwater acoustics, digital signal processing in the field of multimedia communications, and digital content protection and management.