

Improved Watermark Detection for Spread-Spectrum Based Watermarking Using Independent Component Analysis

Hafiz Malik

Dept. of Electrical and Computer
Engineering University of Illinois at
Chicago, Illinois, USA

hmalik@ece.uic.edu

Ashfaq Khokhar

Dept. of Electrical and Computer
Engineering University of Illinois at
Chicago, Illinois, USA

khokhar@ece.uic.edu

Rashid Ansari

Dept. of Electrical and Computer
Engineering University of Illinois at
Chicago, Illinois, USA

ansari@ece.uic.edu

ABSTRACT

This paper presents an efficient blind watermark detection/decoding scheme for spread spectrum (SS) based watermarking, exploiting the fact that in SS-based embedding schemes the embedded watermark and the host signal are mutually independent and obey non-Gaussian distribution. The proposed scheme employs the theory of independent component analysis (ICA) and posed the watermark detection as a blind source separation problem. The proposed ICA-based blind detection/decoding scheme has been simulated using real-world audio clips. The simulation results show that the ICA-based detector can detect and decode watermark with extremely low decoding bit error probability (less than 0.01) against common watermarking attacks and benchmark degradations.

Categories and Subject Descriptors

H.1.1: Systems and Information Theory

General Terms

Algorithms, Design, Performance, Security

Keywords

Spread Spectrum, Independent Component Analysis, Blind Source Separation, Watermarking, Detection, Correlation

1. INTRODUCTION

Technological and socio-economical factors such as the growth of the Internet, the proliferation of low-cost and reliable storage devices, the deployment of seamless broadband networks, the availability of state-of-the-art digital media production and editing technologies, and the development of efficient multimedia data compression schemes have made digital forgeries and unauthorized sharing of digital media a wide spread reality. This form of piracy has subjected the entertainment industry to enormous annual revenue losses. For example, music industry alone claims multi-million illegal music downloads on the Internet every week. It is therefore imperative to develop robust technologies to protect copyrighted digital media from illegal

sharing and tampering. Traditional digital data protection techniques, such as encryption and scrambling, alone cannot provide adequate protection of copyrighted contents, because these technologies are unable to protect digital content once they are decrypted or unscrambled. Digital watermarking technology complements cryptography for protecting digital content even after it is deciphered [1].

Digital watermarking is a process of imperceptibly embedding content protection/authentication information (watermark) into the digital media (the host media). Consequently, watermark detection/decoding process is employed to extract the embedded watermark from the watermarked media. Although the performance expected from a given watermarking system depends on the target application area [1], robust embedding scheme and efficient detection procedure are inherently desired.

The work presented in this paper is focused on investigating efficient watermark detection/decoding schemes. In general, the existing watermark detectors can be classified into two categories: (a) informed detectors that assume that the host signal is available at the detector during watermark detection process, and (b) blind detectors that assume that the host signal is not available at the detector for watermark detection. Similarly, watermark embedding schemes can be classified into two major groups: (a) blind embedding implies that the watermark embedder does not use the host signal information during the watermark embedding process (spread spectrum based watermarking schemes [1, 2, 3, 6] fall in this category), and (b) informed embedding implies that the watermark embedder exploits the properties of the host media during the embedding process (quantization index modulation based watermarking schemes [1, 4, 6] belong to this category).

In order to meet the fidelity requirement of the watermarked signal, the power of the embedded watermark (watermark strength) is generally kept much lower than the host signal power. Blind watermark detection schemes for SS-based watermarking employ statistical characterization of the host signal to develop an optimal or near-optimal watermark detector (in the maximum likelihood (ML) sense) [6]. The salient features of such schemes include increased robustness against interference (i.e. antijamming capability), simplicity, and low computational complexity. The blind watermark detectors for SS-based watermarking perform poorly, particularly in terms of decoding bit error probability at the decoder due to the presence of host-signal interference at the decoder. The nonzero decoding bit error probability at the watermark decoder even in the absence of attack-channel distortion is one of the limitations of the existing blind detectors.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

DRM '05, November 7, 2005, Alexandria, Virginia, USA.
Copyright 2005 ACM 1-59593-230-5/05/0011...\$5.00.

In this paper we develop a novel independent component analysis (ICA) based watermark detection method for SS-based watermarking. In the past ICA based framework has been used for multimedia watermarking [8–13]. However, existing ICA based data-hiding schemes are either not applicable to SS-based watermarking [8–11] or use an informed detector for watermark extraction process [12–13]. For example, Yu et al in [12] have proposed ICA based watermark detector that can be used for SS-based watermarking but their detector uses the watermark and private data during watermark extraction process. Similarly, Sener et al [13] have presented ICA based watermark detector which is also applicable to SS-based watermark detection, but their proposed detector also uses watermark during the detection process, therefore, cannot be used for blind watermark extraction/detection applications.

The main motivation of this paper is to design a blind detector¹ for SS-based watermarking schemes capable of canceling the host-signal interference at the detector, hence improving decoding as well as detection performance. Towards this end, we use the theory of independent component analysis (ICA) by posing watermark detection as a blind source separation problem. The proposed detector assumes that the received watermarked signal is a linear mixture and the underlying independent components (the host signal and the watermark) obey non-Gaussian distributions. The proposed detector falls in the category of blind detectors. We can show that the SS-based watermark estimation/detection problem fits into the blind source separation (BSS) model.

The proposed ICA-based detector first estimates hidden independent components from the received watermarked signal using ICA framework. The estimated components are then used to detect the embedded watermark. Simulation results show that the proposed detector outperforms the correlation based detector (commonly used for watermark detection in SS-based watermarking), against a variety of signal manipulations and degradations applied to the watermarked media. These signal degradations include addition of color and white noise, resampling, requantization, lossy compression, filtering, time- and frequency-scaling, and stirmark benchmark attack [31, 32]. The proposed ICA-based watermark detector is applicable to SS-based watermarking of all media types, i.e. audio, video and images. However, in this paper digital audio is used as the host media for watermark embedding, detection, and performance analysis of the proposed ICA-based detector.

2. BASICS OF SS-BASED WATERMARKING

The spread-spectrum (SS) based watermarking system can be modeled using a classical secure communication model [1], shown in Figure 1. In Figure 1, $\mathbf{s} \in \mathfrak{R}^N$ is a vector containing coefficients of an appropriate transform of the host signal. It is assumed that coefficients, $s[i]: i = 0, 1, \dots, N-1$, are independent and identically distributed (i.i.d.) random variables (r.v.) with zero mean and variance σ_s^2 . A watermark, $\mathbf{w} \in \{\pm 1\}^N$, is generated using (a) an input message b selected from set of

possible embedding messages, and (b) a data embedding secret key K . We assume that the watermark $w[i]$ and the host signal coefficients $s[i]$ are mutually independent. The watermarked signal \mathbf{x} is obtained by adding an amplitude-modulated watermark \mathbf{w} to the host signal \mathbf{s} . In order to meet the fidelity requirement of the watermarked media \mathbf{x} , i.e., to ensure that the embedded watermark is imperceptible, the amplitude-modulated watermark is spectrally shaped according to perceptual mask $\boldsymbol{\alpha}$, estimated based on the human auditory system (HAS) and the host signal \mathbf{s} i.e. $\boldsymbol{\alpha} = f(\mathbf{s}, HAS)$.

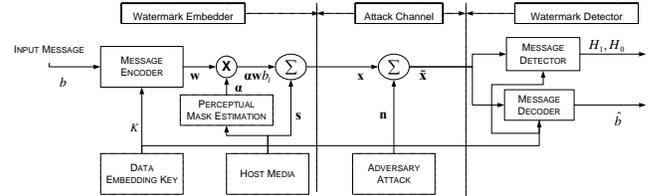


Figure 1: Depiction of a Perceptual Based Data Hiding System with Blind Receiver as a Standard Secure Communication Model

The watermarked signal \mathbf{x} can be expressed as

$$\mathbf{x} = \mathbf{s} + \boldsymbol{\alpha} \cdot \mathbf{w}b \quad (1)$$

where $[\cdot]$ denotes element-wise product of the two vectors.

Embedding distortion generally serves as a fidelity measure for a given watermarking scheme and can be expressed as:

$$\mathbf{d}_e = \mathbf{x} - \mathbf{s} \quad (2)$$

The mean-squared error embedding distortion be expressed as:

$$D_e = \|\mathbf{d}_e\|^2 = \|\mathbf{x} - \mathbf{s}\|^2 = \|\boldsymbol{\alpha} \cdot \mathbf{w}\|^2 = \frac{1}{N} \sum_{i=0}^{N-1} \alpha^2[i] = \sigma_w^2 \quad (3)$$

where $\|\cdot\|$ represents the Euclidian norm.

The signal distortion due to an active adversary attack can be viewed as a channel noise, \mathbf{n} , shown in Figure 1. The resulting watermarked signal, $\tilde{\mathbf{x}}$, is processed for watermark detection and is expressed as:

$$\tilde{\mathbf{x}} = \mathbf{x} + \mathbf{n} \quad (4)$$

The received watermarked signal is used to extract embedded information. The SS-based watermarking schemes use probabilistic characterization of the host data to develop an optimal or near-optimal watermark detector (in ML sense). Pérez-González et al have shown in [6] that the SS-based watermarking schemes are inherently bounded by the host-signal interference at the watermark detector as far as decoding and detection performance at the watermark detector are concerned. The main objective of this paper is to design a watermark detector for SS-based watermarking schemes with improved watermark detection performance by rejecting the host-signal interference at the detector using ICA framework. The fundamentals of ICA theory are briefly outlined in the following section followed by details of the proposed ICA-based detector.

¹ For the rest of the paper we will treat watermark detector as a blind watermark detector unless otherwise stated.

3. INDEPENDENT COMPONENT ANALYSIS

Independent Component Analysis (ICA) is a statistical framework for estimating underlying hidden factors or components of a multivariate statistical data. In the ICA model, data variables are assumed to be linear or nonlinear mixtures of some unknown latent variables, and the mixing system is also unknown [14–16]. Moreover, these hidden variables are assumed to be non-Gaussian and mutually independent. The ICA model can be considered as an extension of the principal component analysis (PCA) and factor analysis [14–16]. In fact, ICA can be treated as non-Gaussian factor analysis, since data is modeled as a linear mixture of some underlying factors. The ICA framework has been used in diverse application scenarios including blind source separation (BSS), feature extraction, telecommunication, and economics [14–16]. In the following we will review only linear ICA framework since that is applicable to the SS-based watermarking model.

In general, the linear ICA model for the noise free as well as noisy observation can be defined as

Noise free ICA model: ICA of a random vector \mathbf{x} consists of estimating the following generative model of the data:

$$\mathbf{x} = \mathbf{A}\mathbf{s} \quad (5)$$

where \mathbf{x} represents the observed m -dimensional random vector, the hidden variables (components) s_i is the vector $\mathbf{s} = (s_1, \dots, s_{n_1})^T$ are assumed statistically independent, and the matrix \mathbf{A} is a $m \times n_1$ mixing matrix.

Noisy ICA model: ICA of a random vector \mathbf{x} consists of estimating the following generative model of the data:

$$\mathbf{x} = \mathbf{A}\mathbf{s} + \mathbf{n} \quad (6)$$

where \mathbf{n} is m -dimensional random noise vector, \mathbf{x} , \mathbf{s} , and \mathbf{A} are as in noise free model.

In this paper, we use noisy ICA generative model to design an ICA-based watermark detector for SS-based watermarking schemes. The proposed ICA-based watermark detector will estimate embedded watermark from the received watermarked signal by reducing the host-signal interference at the watermark detector. In order to understand the basics of ICA, noise free ICA generative model is discussed first. Before estimating the underlying independent components from observed data using ICA framework, the generative model should meet certain conditions to ensure the identifiability of the ICA model. The identifiability constraints, as proposed in [14], for noise free ICA model are outlined as,

1. *Statistical independence:* The hidden (latent) variables/sources are statistically independent.
2. *Non-Gaussianity:* At the most one independent components s_i ; $i = 1, 2 \dots n_1$, is normally distributed.
3. *More sensors than sources:* Number of observed linear mixtures (sensors) m must be greater or equal to the number of independent components (sources) n_1 i.e. $m \geq n_1$.
4. *Full rank mixing matrix:* The rank of the mixing matrix, \mathbf{A} , must be full column rank.

Moreover, \mathbf{x} and \mathbf{s} are assumed to be centered.

Here, third condition (i.e. more sensors than sources) is a relatively loose restriction, as ICA model is identifiable even for

more latent components (sources) than observations (sensors), i.e. $m < n_1$, scenarios [21–24].

The identifiability of the noisy ICA model (Eq. (6)) requires almost the same set of restrictions excluding third, i.e. $m \geq n_1$, and fourth. If we assume that noise is independent from the underlying independent components then the noisy ICA model can be treated as a special case of noise free ICA, that is, degenerate case, e.g. for more sources than sensors scenario i.e. $m < n_1$. For more details on identifiability of noisy ICA model please see [21–25] and references there.

Independence and maximum non-Gaussianity are two fundamental ingredients of the ICA theory. Independence of the underlying components is one of the assumptions to estimate independent components from the linear mixture of the underlying independent components. This should be noted that independence of the underlying components is a much stronger condition than their uncorrelatedness. For example, in case of BSS, there might be many dependent but uncorrelated representations of the observed signals and these uncorrelated but dependent representations of the observed signals can not separate the mixed sources [14], therefore uncorrelatedness itself is insufficient constraint to solve the BSS problem. Actually independence implies nonlinear uncorrelatedness [14], that is, if s_1 and s_2 are two independent components then their any nonlinear transformations say, $\varphi_1(s_1)$ and $\varphi_2(s_2)$, are uncorrelated as well (i.e. their covariance is zero). On the other hand, if s_1 and s_2 are simply uncorrelated then in general, such nonlinear transformations, not necessarily have zero covariance. Thus to perform ICA, a stronger form of decorrelation of the underlying components is required, that is, nonlinear decorrelation. The suitable selection of nonlinearities i.e., $\varphi_1(\cdot)$ and $\varphi_2(\cdot)$ can be achieved using tools like maximum likelihood and mutual information from estimation theory and information theory [7].

Maximum non-Gaussianity is another important requirement of ICA-based hidden components estimation [14–19, 26, 27]. Fourth-order statistics also called kurtosis, k , is generally used as a non-Gaussianity measure of a random variable. The normalized kurtosis of a real r.v. s can be defined as,

$$k(s) = E\{s^4\} / E^2\{s^2\} - 3 \quad (7)$$

here $E\{\cdot\}$ denotes expected value of a r.v.

This is observed that the normal r.v. has zero normalized kurtosis; therefore, kurtosis is a measure of the distance of a r.v. from Gaussianity. In addition, r.v. having positive kurtosis i.e. $k > 0$ are generally called as super-Gaussian, the Laplacian distribution is a typical example of a super-Gaussian r.v.; similarly, r.v. having negative kurtosis value, i.e., $k < 0$ are generally called as sub-Gaussian, e.g., the uniform distribution.

The BSS is one of the most explored applications of the ICA model [14–19, 26, 27]. In BSS, the observed random vector \mathbf{x} corresponds to a realization of an m -dimensional discrete-time signal $\mathbf{x}(t)$, $t = 1, 2, \dots$, and the underlying independent components $s_i(t)$ are called as source signals. The goal is to recover independent sources $s_i(t)$; $i \in n_1$, from the observation \mathbf{x} .

The cocktail party problem is a classical example of BSS where several people are simultaneously speaking in the same room and

the problem is to separate the voices of different speakers (sources), using recordings of several microphones (observations/sensors) in the room.

The ICA model is a suitable framework to solve the cocktail party problem, because the underlying variables are independent in the observed data. In case of SS-based watermarking, the watermarked signal \mathbf{x} at the detector can be treated as an observation which is a linear mixture of the underlying hidden independent components, i.e., the host signal \mathbf{s} and the watermark $\alpha \mathbf{w}$. To this end, the ICA framework can be used to find a linear representation of the underlying components. In other words, BSS using ICA framework is to estimate the demixing matrix \hat{B} based on the observed data \mathbf{x} . The estimated demixing matrix \hat{B} is the inverse (or generalized inverse) of mixing matrix A , i.e.,

$$\hat{B} = \hat{A}^\dagger = (\hat{A}^H \hat{A})^{-1} \hat{A}^H \quad (8).$$

Most of existing BSS schemes using ICA model are based on the information-theoretic framework, for example, Bell et al's [26] ICA scheme is based on the idea of information maximization or "infomax" among the estimated independent components, Comon [15] used higher-order cumulants, Gaeta et al [20] used ML method, and extensions of infomax, higher-order cumulants, ML method etc. can be found in [14, 16, 26, 27].

4. PROPOSED ICA BASED WATERMARK DETECTOR

The proposed ICA-based watermark detector consists of two stages: 1) watermark estimation stage, and, 2) watermark decoding and/or detection stage. The watermark estimation stage estimates watermark $\hat{\mathbf{w}}$ from the received watermarked audio $\tilde{\mathbf{x}}$ using ICA framework, whereas, the watermark decoding and/or detection stage decodes and/or detects the embedded watermark. The block diagram of the proposed watermark detector is given in the Figure 2.

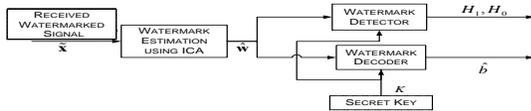


Figure 2: Block Diagram of the Proposed ICA-based Watermark Detector

In general the ICA model for BSS estimates demixing matrix \hat{B} , from the observed data \mathbf{x} . This model is extendable to the watermark estimation problem, assuming identifiability and separability conditions of the ICA model are satisfied. To explore the fact that SS-based watermarking satisfies the identifiability constraints of an ICA model; Let us assume additive SS-based watermark embedding scenario, to this end, the watermarked signal is obtained by a linear combination of the host signal and the watermark. Recalling Eq. (1) with $b = 1$, we have,

$$\mathbf{x} = \mathbf{s} + \alpha \cdot \mathbf{w}$$

Also recall the fact that watermark sequence, \mathbf{w} , and the host signal, \mathbf{s} , are assumed to be mutually independent for all SS-based watermarking schemes. We only need to show that the underlying components are non-Gaussian and the observation satisfies the

full-rank condition. For non-Gaussianity requirement, the real-world audio and image coefficients in DWT domain can be approximated by the Laplacian distribution [5, 6, 26, 32]. If the watermark, \mathbf{w} , is generated based on some non-Gaussian distribution then non-Gaussianity condition of the watermark is also satisfied. The full-rank condition can be relaxed if ICA framework of more sources than sensors [21–25,] is used for watermark estimation. However, for our watermark estimation stage we have used standard ICA model for BSS which requires full-rank constraint along with other three constraints for its identifiability. In order to meet the full-rank condition of the observation, at least one observation independent of the received watermarked audio is required. This is because, in case of SS-based watermarking, the received watermarked signal is a mixture of two independent components, i.e., host signal and watermark (under zero attack-channel distortion). Therefore, in order to meet the full-rank requirement, another observation independent of the received watermarked audio is generated locally, i.e. at the detector, using watermarked signal and an independent pseudo-random noise.

Once identifiability conditions of the standard ICA model are satisfied at the watermark detector, both the noise free ICA model as well as noisy ICA model can be extended to solve the watermark detection problem for SS-based watermarking. Before applying an ICA model for watermark estimation, the following lemma will show that SS-based watermark detection model fits into the ICA model.

Lemma:

SS-based watermark estimation model fits into ICA model.

Proof:

Let \mathbf{s}_1 be the host signal to be watermarked based on additive SS-based watermarking, then the corresponding watermarked signal using Eq. (1) with $b = 1$, can be expressed as

$$\mathbf{x}_1 = \mathbf{s}_1 + \alpha_1 \cdot \mathbf{w} \quad (9)$$

where α_1 is the estimated perceptual mask of the host signal, \mathbf{s}_1 , based on the HAS.

The observed watermarked signal, $\tilde{\mathbf{x}}_1$, at the detector in the presence of additive attack-channel distortion $\mathbf{n}_{attack1}$ can be expressed as

$$\begin{aligned} \tilde{\mathbf{x}}_1 &= \mathbf{x}_1 + \mathbf{n}_{attack1} \\ &= \mathbf{s}_1 + \alpha_1 \cdot \mathbf{w} + \mathbf{n}_{attack1} \end{aligned} \quad (10)$$

In order to meet the full-rank requirement for identifiability of the ICA model, an observation generated which is given as,

$$\begin{aligned} \tilde{\mathbf{x}}_g &= \tilde{\mathbf{x}}_1 + \mathbf{v} \\ &= \mathbf{x}_1 + \mathbf{n}_{attack1} + \mathbf{v} \end{aligned} \quad (11)$$

where \mathbf{v} is the pseudo-random noise, independent of the received watermarked signal, $\tilde{\mathbf{x}}_1$. Here \mathbf{v} can be expressed as,

$$\mathbf{v} = \varepsilon \cdot \mathbf{w} + \mathbf{v}' \quad (12)$$

where $0 \leq \varepsilon \leq 1$

therefore,

$$\begin{aligned}\tilde{\mathbf{x}}_g &= \mathbf{s}_1 + \boldsymbol{\alpha}_1 \cdot \mathbf{w} + \mathbf{n}_{attack1} + \boldsymbol{\varepsilon} \cdot \mathbf{w} + \mathbf{v}' \\ &= \mathbf{s}_1 + (\boldsymbol{\alpha}_1 + \boldsymbol{\varepsilon}) \cdot \mathbf{w} + \mathbf{n}_{attack1} + \mathbf{v}' \\ &= \mathbf{s}_1 + \tilde{\boldsymbol{\alpha}} \cdot \mathbf{w} + \mathbf{n}_g\end{aligned}\quad (13)$$

where $\tilde{\boldsymbol{\alpha}} = \boldsymbol{\alpha}_1 + \boldsymbol{\varepsilon}$, and $\mathbf{n}_g = \mathbf{n}_{attack1} + \mathbf{v}'$

The received watermarked signal and locally-generated observation in matrix form,

$$\begin{bmatrix} \tilde{\mathbf{x}}_1 \\ \tilde{\mathbf{x}}_g \end{bmatrix} = \begin{bmatrix} \mathbf{1} & \boldsymbol{\alpha}_1 \\ \mathbf{1} & \boldsymbol{\alpha} \end{bmatrix} \begin{bmatrix} \mathbf{s}_1 \\ \mathbf{w} \end{bmatrix} + \begin{bmatrix} \mathbf{n}_{attack1} \\ \mathbf{n}_g \end{bmatrix}\quad (14)$$

Let

$$\mathbf{x}'' = \begin{bmatrix} \tilde{\mathbf{x}}_1 \\ \tilde{\mathbf{x}}_g \end{bmatrix}, \mathbf{A}'' = \begin{bmatrix} \mathbf{1} & \boldsymbol{\alpha}_1 \\ \mathbf{1} & \boldsymbol{\alpha} \end{bmatrix}, \mathbf{s}'' = \begin{bmatrix} \mathbf{s}_1 \\ \mathbf{w} \end{bmatrix}, \text{ and } \mathbf{n}'' = \begin{bmatrix} \mathbf{n}_{attack1} \\ \mathbf{n}_g \end{bmatrix}$$

rewriting Eq. (15) we have,

$$\mathbf{x}'' = \mathbf{A}'' \cdot \mathbf{s}'' + \mathbf{n}''\quad (15)$$

therefore, Eq. (15) is equivalent to Eq. (6),

hence, SS-based watermark estimation model fits into the noisy ICA model.

Therefore, an ICA model can be used to solve SS-based watermark estimation problem. \square

In the above proof, we assumed that watermark is embedded into the entire host audio clip (non-segmentation scenario). The ICA model is still extendable to SS-based repeated watermark embedding scenario.

4.1 Watermark Estimation

For watermark estimation, the proposed watermark detector first estimates the watermark-mixing matrix \hat{A}'' which is used to estimate the underlying independent components (i.e., the host signals \mathbf{s} and the watermark \mathbf{w}). An estimate of the watermark-mixing matrix \hat{A}'' is obtained by optimizing some highly nonlinear function also known as contrast function. The pseudo-inverse of the estimated watermark-mixing matrix \hat{A}''^\dagger is applied to the observed mixture to estimate the host signal $\hat{\mathbf{s}}$ and the watermark $\hat{\mathbf{w}}$. However, as noted earlier, in case of blind detectors for SS-based watermarking schemes, watermark estimation using ICA framework is a degenerate case, i.e., $m < n_I$. Hence, just the estimation of watermark-mixing matrix is insufficient to separate the underlying independent components perfectly. In the case of SS-based watermarking, the equation $\mathbf{x} = \hat{A}'' \mathbf{s}$ has an affine set of solutions [28]. A preferred solution in this affine set is generally selected using probabilistic prior model of the independent components. The performance of the proposed ICA-based watermark estimator depends on the separation quality of the separated (estimated) watermark $\hat{\mathbf{w}}$. The separation quality of the separated source is generally measured in terms of *source-to-interference ratio* (*watermark-to-interference ratio* (*WIR*), in case of watermark estimation), *source-to-noise ratio* and *source-to-artifact ratio* (for further

details on these separation quality measures please see [28] and references therein). For performance analysis of the proposed ICA detector, only *WIR* distortion measure is considered here; therefore, the estimated watermark $\hat{\mathbf{w}}$ can be expressed as

$$\hat{\mathbf{w}}[i] = \eta_1 [i] \alpha [i] \mathbf{w}[i] \mathbf{b} + s_{\text{interf}}[i]\quad (16)$$

where η_1 is real constant $0 \leq \eta_1 \leq 1$, and s_{interf} is interference due to the host signal.

Let $s_{\text{interf}} = \eta_2 s$, $0 \leq \eta_2 \leq 1$ then Eq. (16) can be rewritten as,

$$\hat{\mathbf{w}} = \eta_1 \boldsymbol{\alpha} \mathbf{w} \mathbf{b} + \eta_2 \mathbf{s}\quad (17)$$

The relative distortion due to interference in the estimated watermark is defined as,

$$D_{\text{interf}} = (\eta_1 / \eta_2)^2\quad (18)$$

and $WIR = 10 \cdot \log(D_{\text{interf}})$

In general, $D_{\text{interf}} > 1$ for most of the existing BSS schemes that use ICA framework.

Several researchers have proposed elegant BSS algorithms based on ICA model for noisy data [17, 19, 21–25, 27], and these algorithms can be used for watermark estimation. However, we have selected the FastICA for noisy data [17] in our implementations due to its better computational performance and separation quality over existing algorithms [28, 29].

5. SIMULATION RESULTS

In order to evaluate the detection-performance of the proposed ICA-based watermark detector (ICAWD), we applied the conventional normalized correlation based watermark detector (NCWD) [1] to the watermark estimated using ICA framework. The proposed ICAWD can be used to detect watermark for most of the existing SS-based watermark embedding schemes [1–3]. The simulation results presented here are obtained using our frequency selective spread spectrum (FSSS) audio watermarking scheme [3]. For the sake of completeness we briefly discuss embedding and detection details of the FSSS audio watermarking [3].

- **Watermark Embedding:** For the watermark embedding process, we have used same the settings as outlined in [14] except the watermark generation process. The simulation results presented here are based on the watermark, generated using a pseudo-random noise generator obeying some non-Gaussian distribution. The non-Gaussian distribution of the watermark is to satisfy the non-Gaussianity requirement of the ICA model. To improve the watermark detection performance further, watermark is generated using different non-Gaussian distributions depending on encoding symbols. For example, in case of binary embedding, in order to embed symbol ‘0’ watermark w_0 is generated based on non-Gaussian distribution, $f_{0X}(x)$, and to embed symbol ‘1’ watermark w_1 is generated based on a different non-Gaussian distribution, $f_{1X}(x)$. Here $f_0(\cdot)$ and $f_1(\cdot)$ are different

distributions (for the simulation results presented, w_0 is generated based on the Laplacian distribution and w_1 is generated based on the Uniform distribution). In addition, a secret key K_w is used as a ‘seed’ for the pseudo-random noise generator during watermark generation process. The block diagram of the FSSS-based watermark generation and embedding used for simulations is illustrated in Figure 3.

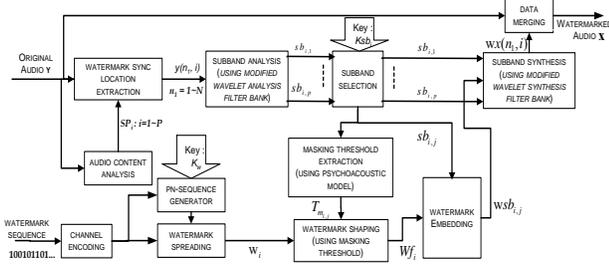


Figure 3: Block Diagram of the FSSS-based Watermark Embedding

• **Watermark Detection:** The proposed ICA-based detector has access to the secret key K only, which is a combination of K_{sb} (subband selection key) and K_w (watermark generation key),

i.e. $K = K_{sb} | K_w$. The watermark detection process for FSSS-based audio watermarking under proposed detection scheme consists of watermark estimation step using ICA framework followed by correlation based detection step, as outlined below,

- **Sync Point Extraction:** The received audio signal is analyzed first to extract the set of *sync points* (SP) [2, 3] in order to combat desynchronization attacks.
- **Segmentation:** An audio frame consisting of N -samples is selected around each SP_i : $i = 1, 2, \dots, M$, where M is the cardinality of set SP.
- **Frame Decomposition:** Each frame is then decomposed into p -subband signals using l -level analysis filter bank discussed in [3].
- **Subband Selection:** A secret key K_{sb_i} , is used to select a subband from lower $p-1$ -subbands of i^{th} frame i.e. $\tilde{s}_{b_{i,j}}$.
- **Watermark Estimation:** The selected subband $\tilde{s}_{b_{i,j}}$ and independently generated pseudo-random Gaussian noise, ν , (defined according to Eq. (12)), are used to generate an independent observation locally (at the watermark detector), in order to meet the full-rank requirement for the identifiability of an ICA model for BSS, i.e.,

$$\tilde{s}_{b_g} = \tilde{s}_{b_{i,j}} + \nu \quad (19)$$

Here, the observation, \mathbf{x} , in matrix form can be expressed as,

$$\mathbf{x} = [\tilde{s}_{b_{i,j}}, \tilde{s}_{b_g}]^T \quad (20)$$

For watermark estimation using ICA framework for BSS, Oja et al’s FastICA for noisy data [17] is used.

• **Information Decoding:** To determine the presence or absence of the watermark in the estimated signals, binary hypothesis testing is employed. For fast and reliable information decoding, normalized correlation between the estimated signals and the key dependent watermarks (i.e. w_0 and w_1) generated at the watermark detector are used. The normalized correlation between \mathbf{w} and estimated signals \mathbf{s}_{ri} is compared against the threshold, Th , to determine the presence or absence of the watermark. The binary hypothesis testing used for binary information decoding is given as,

$$H_1 : \text{if } \max |ncor(\mathbf{s}_{ri}, \mathbf{w}_\xi)| \geq Th, \quad \text{Decode: } \xi$$

$$H_0 : \text{else} \quad \text{no watermark}$$

where $ncor(\cdot, \cdot)$ is the normalized correlation function defined as,

$$ncor(\mathbf{s}_{ri}, \mathbf{w}_\xi) = \frac{\sum_{l=-N}^N s_{ri}(l) w_\xi(n+l)}{\sqrt{\sum_{l=0}^N s_{ri}(l)^2 \cdot \sum_{l=0}^N w_\xi(l)^2}} : \xi = 0, 1 \quad (21)$$

where $\hat{\mathbf{w}}$ is the estimated signals using ICA, Th is the decoding threshold, and $i = 1, 2$.

The block diagram of the proposed ICA-based watermark estimation followed by normalized correlation based detection used for FSSS audio watermark detection process is illustrated in Figure 4.

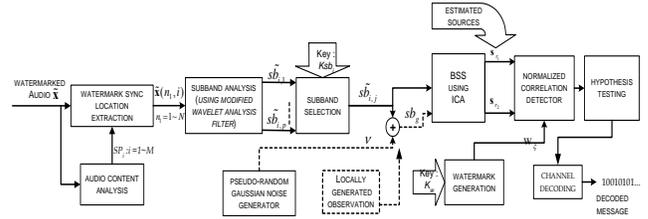


Figure 4: Block Diagram of the proposed ICARD for FSSS-based Audio Watermarking

In order to compare the detection performance of the proposed detector with existing SS-based watermark detection schemes, blind detector based on normalized correlation to the received watermarked audio. Details of the watermark detection process using blind normalized correlation detector for FSSS-based watermarking are outlined in [3]. The block diagram of the blind watermark detector using normalized correlation for watermark detection process used for the FSSS-based audio watermarking scheme is shown in Figure 5.

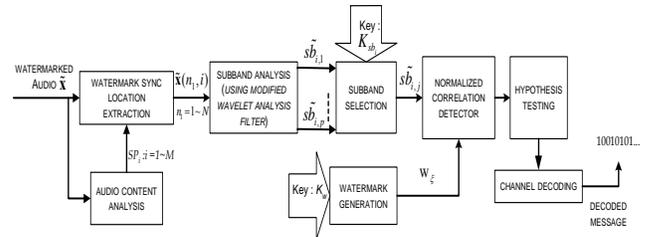


Figure 5: Block Diagram of the Blind NCWD used for FSSS-based Audio Watermarking

5.1 Experimental Results

To evaluate the robustness performance of the proposed watermarking scheme we performed several experimental tests in which the watermarked audio is subjected to commonly encountered degradations. These degradations include addition of white and colored noise, resampling, lossy compression (MPEG Audio compression), filtering, time and frequency-scaling, requantization, and stirmark benchmark attacks for audio. We use decoding bit error probability, Pb_e , at the watermark detector, as a performance measure in this section. The decoding bit error probability, Pb_e , is defined as,

$$Pb_e = 1 - \frac{\text{Number of Bits Correctly Detected}}{\text{Number of Bits Embedded}} \quad (22)$$

The watermark detection results presented in this section are based on an ICAWD (ICA_Detector) and a NCWD (Cor_Detector) [3]. The detection performance, in terms of Pb_e assessed in the presence of different attacks, reveals that an ICAWD outperforms the NCWD due to its better host-signal interference cancellation capability. The penalty for this improved performance is higher computational cost ICAWD.

The simulation results presented here are based on five audio clips belonging to four different music categories (rock/pop, melodic, vocal, and Indian semi-classical). In Table 1, the name, type, and duration of the selected audio clips are listed.

The detection results presented in this section assume, (a) no channel coding, (b) same watermark is embedded into two frames around consecutive SPs, that is, if watermark w is embedded in j^{th} subband of i^{th} frame then w is also embedded in j^{th} subband of $(i+1)^{\text{th}}$ frame, (c) wavelet decomposition level, l is set to 5, and (d) hypothesis testing threshold, Th , is set to the value of 0.15 which corresponds to a false positive rate of 3.5×10^{-4} .

- **Addition of White Noise:** White Gaussian noise with power from zero to 200 % of the power of the audio signal is added to the corresponding watermarked audio clips. The detection performance, Pb_e , of ICAWD and NCWD for different SNR values is plotted in Figure 6. The ICAWD clearly outperforms the NCWD and the improvement may be attributed to its better host-signal interference cancellation ability. Figure 6 also shows that for SS-based watermarking very low decoding bit error probability is possible even in the presence of noise with power 60 - 70 % of the power of the audio signal, given ICA-based detector is used of watermark detection process.
- **Addition of Colored Noise:** To simulate an attack with colored noise, white Gaussian noise is spectrally shaped according to the estimated masking threshold. The masking threshold of watermarked audio clip is estimated based on the HAS model [1, 7]. The just audible colored noise is then added to the watermarked audio signal. The detection performance, Pb_e , for ICAWD and NCWD, for each watermarked audio clip, is presented in Figure 8. This is observed that the NCWD performs poorly, this is due to the increase in interference level (colored noise); as the colored noise is generated with a process profile almost identically to that of the watermark. The colored noise acts as a second watermark, interfering with the watermark to be detected. On the other hand, ICAWD is effective in handling such attacks due to its interference cancellation ability and Figure 8 reflects this fact.

- **Rescaling:** Rescaling attacks include time- and frequency-scaling. Time-scaling attacks can be used to desynchronize a watermark detector, for SS-based watermarking systems. To test the robustness of our proposed detector against time-scaling attacks, the watermarked audio signals are time-scaled by a factor of $ts = \pm 1\%$. The detection performance for ICA-based detector along with correlation based detector is illustrated in Figure 9. Similarly, the frequency-scaling attacks are generally used to adversely affect the detection performance of frequency domain watermarking schemes. To mount frequency scaling attack, the watermarked audio signals are frequency scaled by a factor of $fs = \pm 1\%$. The detection performance in terms of Pb_e of the ICAWD and NCWD, against frequency-scaling attack, is presented in Figure 10. The better performance of ICAWD is evident in Figures 9 and 10.
- **Resampling:** To simulate resampling attack, watermarked audio signal is down-sampled at a sampling rate of f_s / r_f (where r_f is the down-sampling factor) and then up-sampled back to f_s . The watermark detection is applied to this resampled watermarked audio clip. The detection performance, Pb_e , for each audio clip for $r_f = 2, 3, 4, \dots, 10$, is illustrated in Figure 11. Figure 11 shows that ICAWD performs reasonably well against resampling attacks up to $r_f = 5$ for each watermarked audio clip.
- **Requantization:** To simulate a requantization attack, each watermarked audio clip is quantized using 8-bit quantization level. The detection performance after a requantization attack is illustrated in Figure 7.
- **Lossy Compression:** Lossy audio compression, e.g., MP3, is generally applied to the digital audio for multimedia applications like transmission and storage to reduce the bit rate. To test the survivability of the watermark, the MP3 encoding/decoding is applied to the watermarked audio using ISO/MPEG-1 Audio Layer III [7] coder for different bit rates i.e. 32, 64, 96, 112, 128, 192, 256, and 320 k bits/s (kbps). The detection performance after lossy compression attack for these bit rates for ICAWD and NCWD is illustrated in Figure 12 and 13 respectively.
- **Filtering:** To test the robustness of the proposed watermark detector against filtering attacks, the watermarked audio signal is subjected to lowpass filtering, highpass filtering, and bandpass filtering attacks. To test the performance for low-pass filtering the watermarked audio signal is passed through a low-pass filter with $f_c = 5$ kHz and 12 dB/Octave roll-off. The detection performance after lowpass filtering attack is illustrated in figure 14a. Similarly, for highpass filtering attack, the watermarked audio signal is passed through the first-order highpass filter with $f_c = 100$ Hz. The detection performance against highpass filtering attack, for all watermarked audio clips, is presented in Figure 14b. And 1st- order bandpass filtering attack, with cut-off frequencies $f_{c_low} = 50$ Hz, and $f_{c_up} = 4.5$ kHz, are presented in Figure 14c.
- **Stirmark Audio Benchmark:** We have also tested the performance of ICAWD and NCWD for FSSS-based watermarking against stirmark audio benchmark [31] attacks. Watermark detection results on the watermarked audio clips degraded with the stirmark audio benchmark attack [32] are listed in Table II. These detection results are for the attack parameters from the version of the tool available online [32]. The detection performance results presented in Table II are averaged over all audio clips listed in Table 1.

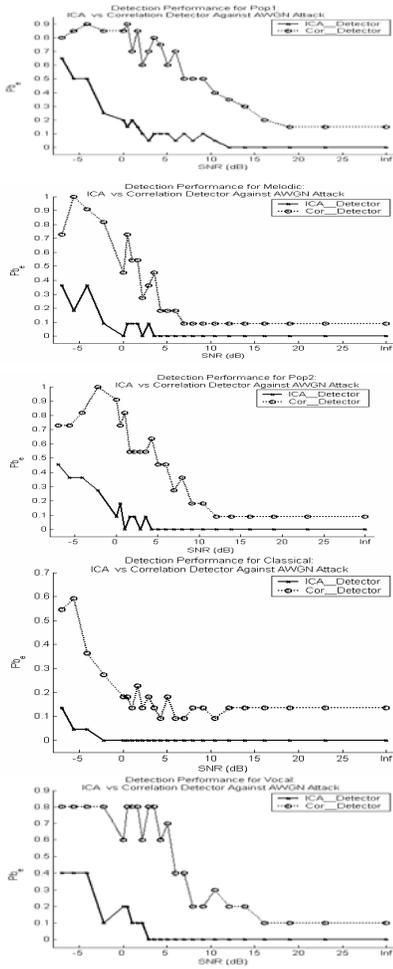


Figure 6: Decoding Performance, P_{b_e} , for an ICAWD and NCWD against AWGN attack for different SNR (dB) values.

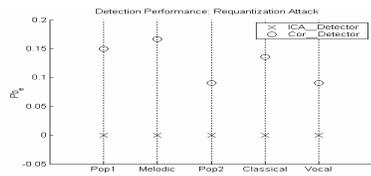


Figure 7: Decoding Performance, P_{b_e} , for Requantization attack applied to each Watermarked Audio Clip.

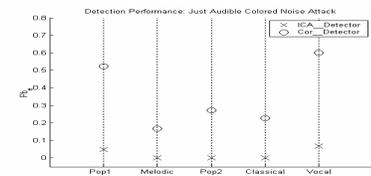


Figure 8: Decoding Performance, P_{b_e} , against Just Audible Colored Noise attack applied to each Watermarked Audio Clip.

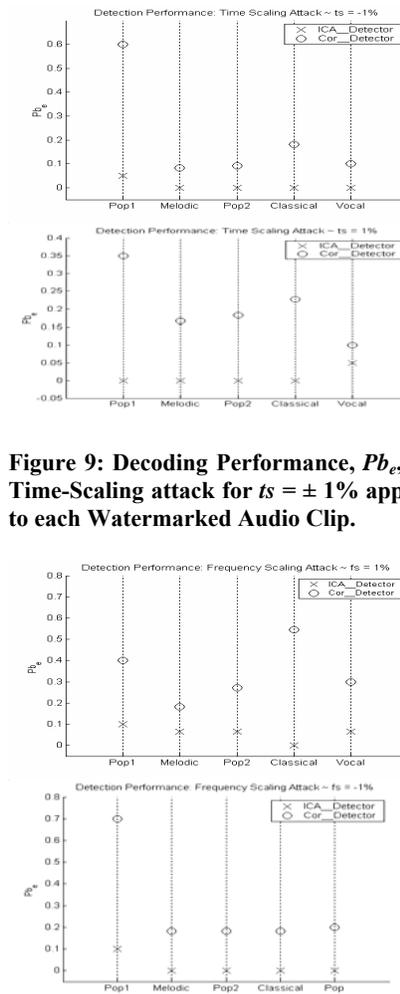


Figure 9: Decoding Performance, P_{b_e} , for Time-Scaling attack for $ts = \pm 1\%$ applied to each Watermarked Audio Clip.

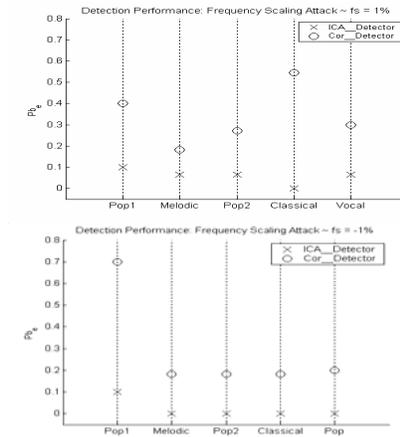


Figure 10: Decoding Performance, P_{b_e} , for Frequency-Scaling attack for fs Values $\pm 1\%$ applied to each Watermarked Audio Clip.

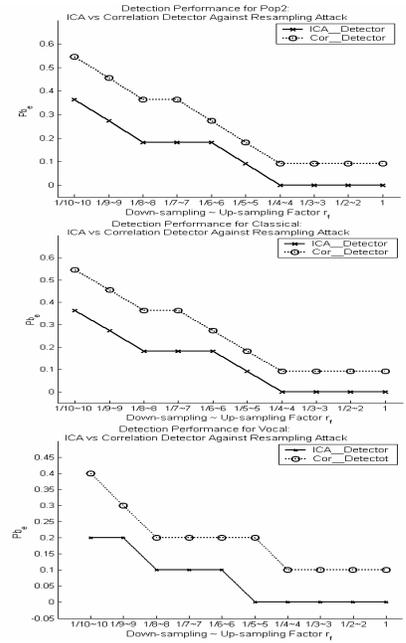


Figure 11: Decoding Performance, P_{b_e} , of each Audio Clip, against Resampling attack for different values of Resampling Factor, r_f .

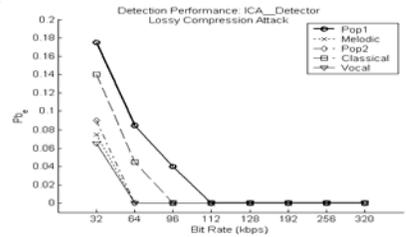


Figure 12: Decoding Performance, P_{b_e} , against Lossy Compression attack for different bits rates using ICAWD applied to each Watermarked Audio Clip.

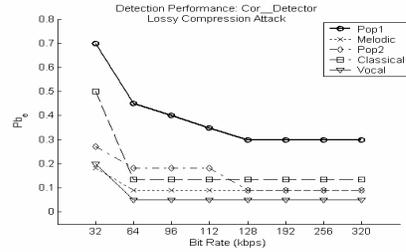
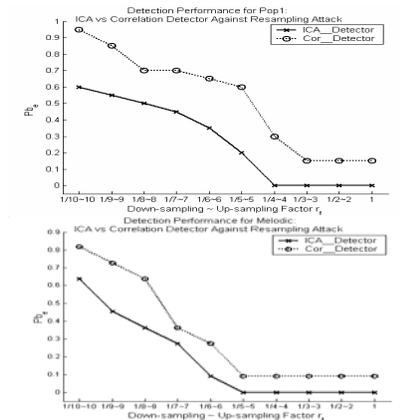


Figure 13: Decoding Performance, P_{b_e} , against Lossy Compression attack for different bits rates using NCWD applied to each Watermarked Audio Clip.

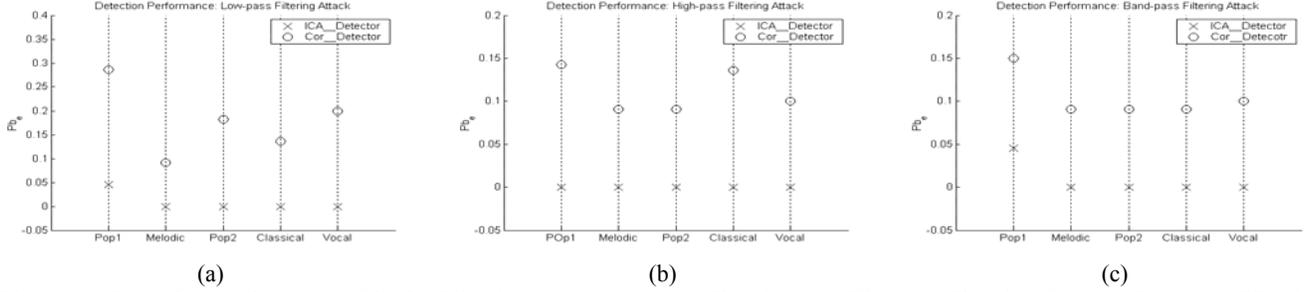


Figure 14: Decoding Performance, P_{b_e} , for Filtering attack, Lowpass Filtering (a), Highpass Filtering (b), and Bandpass Filtering (c), applied to each Watermarked Audio Clip,

TABLE I
SELECTED AUDIO CLIPS

	Singer Name /Song Title	Type	Duration (Sec)
1	Backstreet Boys, I Want It That Way ...	Pop, (Pop1)	22
2	Lata Mangeshkar, Kuch Na Kaho ...	Melodic, (Melodic)	15
3	Asha Bhosle, and Richa Sharma, Kahin Aag Laga ...	Pop, (Pop2)	10
4	Nusrat F. A. Khan, Afreen Afreen ...	Indian Semi-Classic, (Classical)	20
5	Suzanne Vega, Tom's diner	Female Vocal, (Vocal)	5

Here, all Audio Clips are Mono, Sampled at 44.1 kHz, and 16 Bit Resolution.

TABLE II
AVERAGE DETECTION PERFORMANCE RESULTS ON WATERMARKED AUDIO CLIPS ATTACKED WITH THE STIRMARK AUDIO BENCHMARK

Stirmark Attack	$P_{b_e_ICA}$	$P_{b_e_Cor}$
<i>addbrumm_100</i>	0	0.2258
<i>addbrumm_1100</i>	0	0.2258
<i>addbrumm_2100</i>	0	0.2258
<i>addbrumm_3100</i>	0	0.2258
<i>addbrumm_4100</i>	0	0.2581
<i>addbrumm_5100</i>	0	0.2581
<i>addbrumm_6100</i>	0	0.2581
<i>addbrumm_7100</i>	0.0323	0.2903
<i>addbrumm_8100</i>	0.0323	0.3226
<i>Addbrumm_9100</i>	0.0323	0.3226
<i>Addbrumm_10100</i>	0.0646	0.3548
<i>addnoise_100</i>	0	0.2258
<i>addnoise_300</i>	0	0.2258
<i>addnoise_500</i>	0	0.2258

Stirmark Attack	$P_{b_e_ICA}$	$P_{b_e_Cor}$
<i>addnoise_700</i>	0	0.2258
<i>addnoise_900</i>	0	0.2258
<i>Addsinus</i>	0	0.2258
<i>Amplify</i>	0	0.2258
<i>Compressor</i>	0	0.2581
<i>Dynnoise</i>	0	0.2581
<i>Echo</i>	0.0323	0.3548
<i>exchange</i>	0	0.2258
<i>extrastereo_30</i>	0	0.2258
<i>extrastereo_50</i>	0	0.2258
<i>extrastereo_70</i>	0	0.2258
<i>fft_hlpass</i>	0.0323	0.2258
<i>fft_invert</i>	0	0.2258
<i>fft_real_reverse</i>	0	0.2258

Stirmark Attack	$P_{b_e_ICA}$	$P_{b_e_Cor}$
<i>fft_stat1</i>	0.1931	0.4839
<i>fft_test</i>	0.1931	0.4839
<i>Flippsample</i>	0.1613	0.4839
<i>Invert</i>	0	0.2258
<i>Lsbzero</i>	0	0.2258
<i>Normalize</i>	0	0.2258
<i>rc_highpass</i>	0.0323	0.2258
<i>rc_lowpass</i>	0	0.2258
<i>Smooth</i>	0	0.2258
<i>smooth2</i>	0	0.2581
<i>stat1</i>	0	0.2258
<i>stat2</i>	0	0.2258
<i>Zerocross</i>	0	0.2258
<i>Zeroremove</i>	0.0323	0.2258

6. CONCLUSION

An improved watermark detector for SS-based watermarking is presented in this paper. The proposed watermark detector is capable of canceling the host-signal interference at the watermark detector using ICA framework. The proposed ICA-based detector can be used for SS-based watermarking for all types of

multimedia data, e.g., audio, video, images, etc. Simulation results for real-world data show that the proposed ICA-based detector performs much better than the conventional normalized correlation based detector. Moreover, the detection performance of the proposed detector can be improved further by employing channel coding. Theoretical results of the proposed ICAWD are available in [33]. Currently we are investigating the performance

of the proposed ICAWD for SS-based image and video watermarking. We are also looking forward to use the proposed ICA-based detector for multimedia fingerprinting for secure multimedia distribution on the Web and multimedia usage tracking applications.

7. References

- [1] I. J. Cox, M. L. Miller, and J. A. Bloom, *Digital Watermarking*, Morgan Kaufmann, 2001.
- [2] C.-P. Wu, P.-C. Su, and C.-C. J. Kuo, "Robust Audio Watermarking for Copyright Protection," *SPIE's 44th AMASPAAI*, 1999.
- [3] H. Malik, A. Khokhar, and R. Ansari, "Robust Audio Watermarking using Frequency Selective Spread Spectrum Theory," *Proc. ICASSP'04*, Canada, May 2004.
- [4] B. Chen and G. W. Wornell, "Quantization index modulation: A class of provably good methods for digital watermarking and information embedding," *IEEE Trans. on Information Theory*, vol. 47(4), pp. 1423-1443, May 2001.
- [5] S. Mallat., "A theory for multiresolution signal decomposition, the wavelet representation", *IEEE Trans. Pattern Ana. and Machine Intell.*, vol. 2(7), 1989.
- [6] F. Pérez-González, F. Balado, and J. R. Hernández, "Performance analysis of existing and new methods for data hiding with known-host information in additive channels," *IEEE Trans. on Signal Processing*, 51(4):960-980, April 2003.
- [7] P. Noll, "MPEG Digital Audio Coding," *IEEE Sig. Proc. Mag.* vol. 14(5), pp. 59-81, Sep 1997.
- [8] S. Noel, and H. Szu, "Multimedia Authenticity with ICA Watermarks," *Wavelet Applications VII, SPIE Proc.* vol. 4056, pp. 175-184. April, 2000.
- [9] F. Serrano, and J. Fuentes, "Independent Component Analysis Applied to Digital Image Watermarking," *Proc. ICASSP'01*, 2001.
- [10] B. Toch, D. Lowe, and D. Saad, "Watermarking of Audio Signals using Independent Component Analysis," *Proc. 3rd Int. Conf. WEB Delivering of Music*, 2003.
- [11] S. Bounkong, B. Toch, D. Saad, and D. Lowe, "ICA for Watermarking Digital Images," *J. Machine Learning Research 1*, pp. 1-25, 2002.
- [12] D. Yu, F. Sattar, and K. Ma, "Watermark Detection and Extraction using Independent Component Analysis," *EURASIP J. Applied Signal Processing*, pp. 92-104, January 2002.
- [13] S. Sener, and B. Günsel, "Blind Audio Watermark Decoding using Independent Component Analysis," *Proc. 17th Int. Conf. Patt. Reco. (ICPR'04)*, vol. 2 pp. 875-878, 2004.
- [14] A. Hyvärinen, J. Karhunen, and E. Oja, *Independent Component Analysis*, John Wiley & Sons, 2001.
- [15] P. Comon, "Independent Component Analysis, A New Concept?," *Signal Processing, Elsevier*, 36(3):287-314, 1994.
- [16] A. Hyvärinen, and E. Oja, "Independent Component Analysis: Algorithms and Applications" *Neural Networks*, 13(4-5):411-430, 2000.
- [17] A. Hyvärinen, "Fast independent component analysis with noisy data using gaussian moments," *Proc. ISCS'99*, 1999.
- [18] J.F. Cardoso, "High-order contrasts for independent component analysis" *Nural Computation*, vol. 11(1), pp. 157-192, 1999.
- [19] A. Hyvärinen, "Independent Component Analysis in the Presence of Gaussian Noise by Maximizing Joint Likelihood" *Neurocomputing*, 22:49-67, 1998.
- [20] M. Gaeta, and J.-L. Lacoume, "Source separation without prior knowledge: the maximum likelihood solution," *Proc. EUSIPCO'90*, pp. 621-624, 1990.
- [21] A. Cichocki, S. Douglas, and S. Amari, "Robust techniques for independent component analysis (ICA) with noisy data," *Neurocomputing*, vol. 22, pp. 113-129, 1998.
- [22] P. Pajunen., "Blind Separation of Binary Sources with Less Sensors than Sources," *Proc. Int. Conf. on Neural Networks (ICNN-97)*, pp. 1994-1997, 1997.
- [23] M. Zibulevsky, and Y.Y. Zeevi, "Extraction of a single source from multichannel data using sparse decomposition", *Neurocomputing*, 49, pp 163-173, 2002.
- [24] Y. Li, A. Cichocki, and S. Amari, "Analysis of sparse representation and blind source separation" *Neural Computation*, vol. 16(6), 2004.
- [25] P. Comon, "Blind Identification in Presence of Noise," *Proc. EUSIPCO'92*, pp. 835-838, 1992.
- [26] A. J. Bell, and T.J. Sejnowski, "An information maximisation approach to blind separation and blind deconvolution," *Neural Computation*, 7(6), pp. 1129-1159, 1995.
- [27] É. Moulines, J-F. Cardoso, and E. Gassiat, "Maximum likelihood for blind separation and deconvolution of noisy signals using mixture models," *Proc. ICASSP'97*, pp. 3617-20, 1997.
- [28] R. Gribonval, L. Benaroya, E. Vincent, and C. Févotte, "Proposals for Performance Measurement in Source Separation," *4th Int. Sym. ICA & BSS (ICA 2003)*, April, 2003.
- [29] Y. Li, D. Powers, and J. Peach, "Comparison of Blind Source Separation Algorithms," *Advances in Neural Networks and Applications*, N. Mastorakis (Ed.), WSES, pp. 18 - 21, 2000
- [30] M. Steinebach, A. Lang, J. Dittmann, and F. A. P. Pritcolas, "StirMark Benchmark: Audio Watermarking Attacks based on Lossy Compression," *Proc. SPIE Security Watermarking Multimedia*, vol. 4675, pp. 79 - 90, 2002.
- [31] "StirMark Benchmark for Audio", <http://amsl-smb.cs.uni-magdeburg.de/smfa/main.php>
- [32] http://multimedia.ece.uic.edu/~hafiz/ICA_WM.html