Watermarking of audio signals using Independent Component Analysis

Borémi Toch, David Lowe, David Saad
Neural Computing Research Group
Aston University
Birmingham, B4 7ET, United Kingdom
tochb,d.lowe,d.saad@aston.ac.uk

Abstract

A novel approach to watermarking of audio signals using Independent Component Analysis (ICA) is proposed. It exploits the statistical independence of components obtained by practical ICA algorithms to provide a robust watermarking scheme with high information rate and low distortion. Numerical simulations have been performed on audio signals, showing good robustness of the watermark against common attacks with unnoticeable distortion, even for high information rates. An important aspect of the method is its domain independence: it can be used to hide information in other types of data, with minor technical adaptations.

1 Introduction

Most current methods aimed at controlling the distribution of digital information rely on file formats, and are usually implemented by incorporating additional data. Watermarking, however, uses the digital information as a cover signal for hiding messages. The embedded message should be difficult to remove and induce unnoticeable distortion.

Watermarking has numerous applications. The most obvious ones are copyright protection and transaction tracking, but authentication, copy control, device control and broadcast monitoring can also be achieved.

The method proposed in this paper is based on Independent Component Analysis (ICA). It promises higher information rates for a given fidelity and robustness against a range of attacks.

A general model for watermarking will first be outlined, followed by the proposed ICA based watermarking system. Technical details of the method will then be provided. Finally, audio watermarking simulation results will be presented and discussed.

2 Watermarking Model

A watermarking system can be considered as a communication channel: a message is sent from the watermark encoder to the decoder through some communication channel. Let us denote the original cover text or signal by the real vector $\boldsymbol{x}$. The binary message vector to be embedded into $\boldsymbol{x}$ is denoted by $\boldsymbol{m}$, using a watermarking encoder E. Additional information $\boldsymbol{z}$ may be required, such as a private key. The watermarked signal $\hat{x} = E(\boldsymbol{x}, \boldsymbol{m}, z)$ is then conveyed; various attacks may occur during transmission, such as corruption by noise, lossy compression, or malicious attacks aimed at removing the watermark. These attacks are denoted by D so that $\tilde{\boldsymbol{x}} = D(\hat{\boldsymbol{x}})$. Finally, the receiver estimates the message from the attacked cover signal: $\tilde{\boldsymbol{m}} = R(\tilde{\boldsymbol{x}}, \boldsymbol{z})$.

The watermarking system is subject to several requirements. The similarity between the original and watermarked signal is called *fidelity*, and depends on the selected distortion measure.

A watermark should be retrievable from the cover signal even if the latter has been been subject to distortion; this ability is called *robustness*. Attacks can be sorted into two types; malicious, aimed at removing the watermark, and non malicious, such as common signal processing or lossy compression. Robustness is assessed with respect to a given set of attacks.

The *information rate* of a watermarking system is the amount of information embedded. In the case of lossless transmission, the information rate is bounded by the channel capacity, subject to attack and fidelity constraints.

As can be seen from the previous definitions, these three desired but contradictory requirements are tightly linked; for instance, if a high information rate coupled with a high robustness are required, then the fidelity will be compromised. A good watermarking system is therefore application dependent.

0-7695-1935-0/03 $17.00 © 2003 IEEE

3 ICA as a watermarking tool

Most watermarking methods usually apply a transformation to the cover signal, such as a Discrete Fourier Transform (DFT) or Discrete Cosine Transform (DCT). The watermark is then embedded in the transform domain.

Noise-free ICA [2] is a principled method for evaluating statistically independent latent variables $\boldsymbol{s}$ from observations $\boldsymbol{x}$, given that $\boldsymbol{x} = A\boldsymbol{s}$. A is a constant $l \times n$ mixing matrix for simplicity and ease of implementation; l and n are the latent variables and observation vectors, respectively. The corresponding $n \times l$ demixing matrix W admits $\boldsymbol{s} = W\boldsymbol{x}$ and $WA = I_n$, where I_n is the identity matrix in an n-dimensional space.

ICA is traditionally used for Blind Source Separation (BSS); here a mono audio signal is considered as a time series $\mu(t)$ whose values fall in the interval $[-1, 1]$, with $t \in [0, T]$. The samples $x(i)$ are built as delay vectors of size l: $x(i) = (\mu(il), \ldots, \mu(il + l - 1))$, with $i \in [1, T/l]$. The aim is to find a convenient linear decomposition of the signal. Unlike the BSS problem, the meaning of each source has no interest in itself.

The motivation for selecting ICA as an embedding space is the statistical independence of the resulting sources. ICA can be performed using different methods, but they are proved to be almost equivalent, as shown in [5, 2]. One of these methods is based on mutual information minimisation. It is well adapted to the proposed use of ICA, since the sources are not supposed to be strictly independent. Instead, it provides sources which are as independent as possible.

From a watermarking point of view, statistical independence intuitively results in minimal interference between channels, so that watermarking one channel will not cross-interfere with others [1]. Moreover, having an estimate of one source will provide no information about the others, which is a disadvantage for a potential attacker. Independent sources have been proved to induce a higher watermarking capacity in special cases where the original cover signal is known at the decoding stage [6].

The encoding algorithm is given by

$$\hat{\boldsymbol{x}} = A\,E(W\boldsymbol{x}, m) + (\ker(W) \cap \boldsymbol{x}) \;, \tag{1}$$

where E is the encoder, $\boldsymbol{m}$ the binary message to encode and $\ker(W)$ the kernel of the linear transformation W.

The demixing matrix W is communicated to the decoder, and acts as a secret key.

4 Generating the mixing and demixing matrices

In the proposed scheme, the ICA matrices are not generated from the signal to be watermarked, but from a corresponding training set. Thus, the sources used for the signal to be watermarked will not be as independent as possible. But this approach has some advantages.

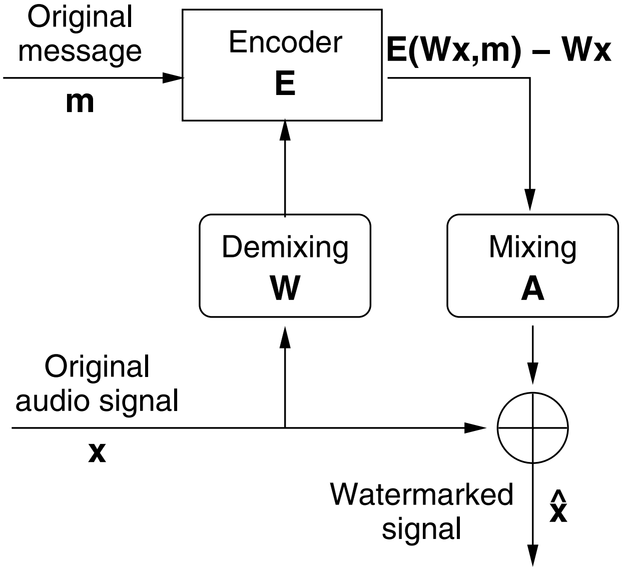

Figure 1. Encoding using ICA decomposition.

The security of the secret keys is assured by the sensitivity of the ICA procedure to the training set. As an example, a test signal $\boldsymbol{x}$ has been watermarked using a demixing matrix W generated from a training set, using the popular FastICA package [3]. The original sources $\boldsymbol{s} = W\boldsymbol{x}$ have been modified to get the watermarked sources $\hat{\boldsymbol{s}} = W\hat{\boldsymbol{x}}$. Then, an attacker who estimates W and $\boldsymbol{x}$ using some ICA procedure, will arrive at estimated demixing matrix $W' \neq W$ and estimated sources $\boldsymbol{s}' = W'\boldsymbol{x}$.

For a given estimated source ${s'}_i$, the most similar original source s_i^{sim} is identified by the corresponding maximum of the mutual information:

$$I(s_i^{sim}; {s'}_i) = \max_j \{I(s_j; {s'}_i)\} \;. \tag{2}$$

These maxima have been plotted in figure 2 as the solid line. The dashed line above is the mutual information between the original components and the decoded components: $I(\hat{s}_i, s_i)$; it is almost equal to the entropy $H(s_i)$ of the original sources.

The large distance between the two curves indicates that the estimated sources ${s'}_i$ are far from the correct sources s_i. Thus, no malicious attack can rely on ICA estimation from the target music file to find the secret key W or to retrieve the unwatermarked cover signal.

Another advantage given by precomputed fixed matrices is a much lower computational cost. Instead of performing an ICA on every signal to be watermarked, the only operation to be carried out is a simple matrix multiplication.

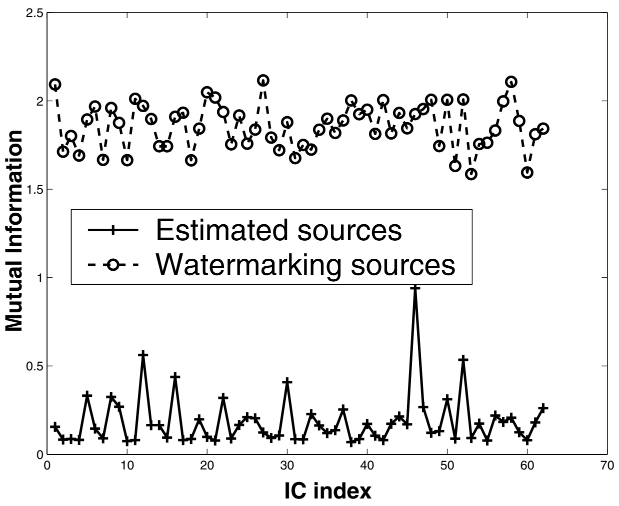

Figure 2. Watermarking process estimation from the watermarked signal.

The training set consists of music fragments of the same genre. The notion of similarity used is very intuitive and is not rigorously justified. For the moment, training sets are composed of fragments of the same record, but different from the signal to be watermarked.

In order to see how good the similarity assumption is, three different examples of different music style have been used: a piece of chamber music, a concerto for piano and tracks from the same rock album. An ICA procedure has then been performed on each training set, in order to get the demixing matrices. Using the previously defined notation, W is the appropriate demixing matrix, coming from a training set similar to the test signal to be watermarked. The other two training sets give two other matrices W_{o1} and W_{o2}. Performing ICA on the test signal gives W', the estimated demixing matrix. Presented for comparison purposes, the matrix performing the DCT transformation is denoted by DCT. Then, these matrices are used to decompose the test signal.

The performance of the different demixing matrices is measured by the pairwise mutual informations of the resulting decomposition: if $\{y_i\}$ is a decomposition, the pairwise mutual informations are defined by $\{I(y_i, y_j)\ ,\ i \neq j\}$.

The means and variances of these mutual informations for each decomposition have been plotted in figure 3. Since ICA minimises the mutual information, the bar corresponding to W' is the lowest, denoted by a dashed line. The appropriate W matrix outperforms the two other demixing matrices W_{o1} and W_{o2} derived from other training sets. The DCT gives the worst decomposition. Thus, the assumption of similarity between the training set and the test signal is valid for these examples.

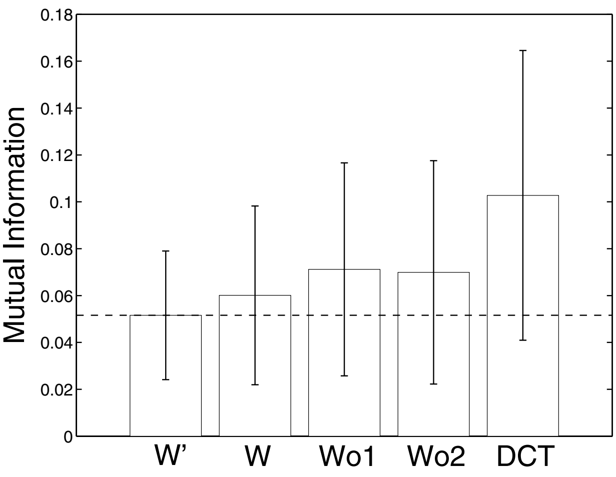

Figure 3. Pairwise mutual informations: mean and standard deviation.

5 Watermarking performance

A simple implementation of ICA watermarking will be presented. The data used are mono signals of audio Compact Disc quality (sampling frequency of 44.1 kHz and 16 bit precision, equivalent to a 90 dB dynamic range). The embedding process is achieved by quantisation of randomly chosen ICA components. For decoding, only the demixing matrix, the watermarking distortion level and the random generator seed are available.

The information rate is fixed by the block size used. In the following examples, the blocks are 512 samples long, which correspond to about 11.6 ms of signal. The encoding is performed using Quantisation Index Modulation (for details see [1] and references within), with one embedded information bit per block, resulting in 86.13 bit/s. Thus, the information rate is $1.2207.10^{-4}$ embedded bits per cover signal bit. For comparison purposes, the information rate of the watermarking system used for DVD audio is $1.2153.10^{-6}$ bits per cover signal bit, corresponding to 2.8 bit/s.

The measure used in the algorithm to quantify the watermark induced distortion is the Peak Signal-to-Noise Ratio (PSNR), much simpler to implement than the feedback algorithms used for psychoacoustic modelling. But since for music PSNR distortion is not very meaningful, an equivalence between PSNR and psychoacoustic distortion has been established, using the Perceptual Evaluation of Audio Quality [4]. For the considered test files, almost perfect fidelity is achieved with a PSNR of about 35 dB and above. At this distortion level, no difference can be heard on consumer level playback devices. At 25 dB, very good quality is achieved for portable audio.

Two fidelity levels are examined: 25 dB and 35 dB. The first attack is jamming. A white noise pattern with random phase is added to the attacked signal, with different amplitudes in order to match several distortion levels. The results are plotted in figure 4.

Both watermarks present less than 5% decoding error rate for a noise 10 dB stronger than the watermark-induced distortion. A noise level of 20 dB is severe, and degrades significantly the cover signal.

The popular Mpeg 1 layer 3 lossy compression scheme has been used as the second attack. Using a psychoacoustic model, it maximises quality for a given average bit rate value. As shown in Figure 5, even under very high and noticeable compression, both watermarks are identified with a low error rate. For the 25 dB watermark, the decoding is almost perfect, making it highly suitable for portable audio purposes, where small size is a critical issue.

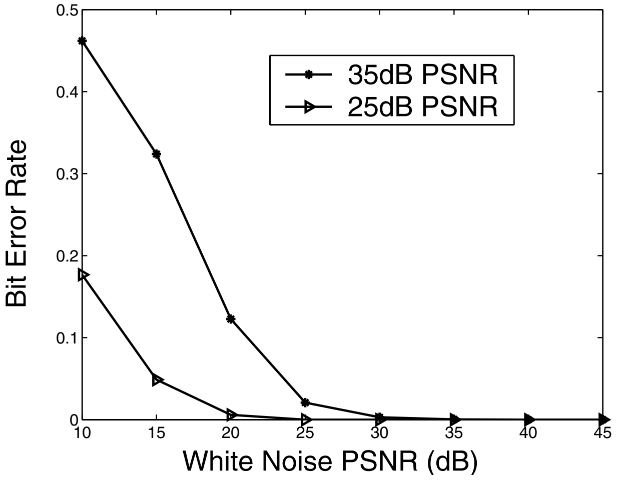

Figure 4. Jamming error rates.

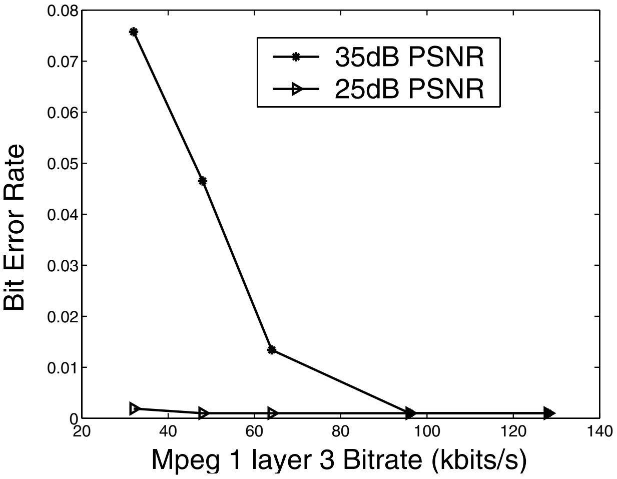

Figure 5. Lossy compression error rates.

6 Conclusion

ICA has been shown to have several interesting properties for watermarking. From an information-theoretic point of view, independence has been proved to be optimal in some cases [6]. If the demixed sources are considered as communication channels, independence intuitively lowers cross-channel interference. The proposed watermarking method relies on ICA matrices generated from training sets. It offers protection against malicious attacks, since estimating the secret keys from the watermarked file is difficult, while allowing a low mutual information between sources (high independence). This watermarking method can be applied to any domain, since it does not directly depend on the nature of the signal [1].

The implemented simple watermarking system based on ICA and quantisation shows good performance in a typical intellectual property rights protection situation. It causes low perceptible distortion while resisting strong non-malicious attacks. The very high information rate can further improve robustness by means of error correcting codes. The robustness against high compression rates can also be used in steganographic applications to provide additional information about the cover signal, such as lyrics, title or artist name.

References

[1] S. Bounkong, B. Toch, D. Saad, and D. Lowe. ICA for watermarking. To be published in Journal of Machine Learning Research: special issue on ICA, 2003.

[2] A. Hyvärinen. Survey on independent component analysis. *Neural Computing Surveys*, 2:94–128, 1999.

[3] A. Hyvärinen, J. Karhunen, and E. Oja. *Independent Component Analysis*. John Wiley & Sons, Inc., 2001.

[4] P. Kabal. An examination and interpretation of ITU-R BS.1387: perceptual evaluation of audio quality. Technical report, Department of Electrical and Computer Engineering, Mc Gill University, 2002. http://www.tsp.ece.mcgill.ca.

[5] T. Lee, M. Girolami, A. Bell, and T. Sejnowski. A unifying information-theoretic framework for independent component analysis. *Computers and Mathematics with Application*, 11(2):1–21, 2000.

[6] P. Moulin and J. O'Sullivan. Information-theoretic analysis of information hiding. *IEEE Transactions on Information Theory*, 49(3):563–593, Mar. 2003.