Audio Watermarking & Its Applications

[1]Sanjay Kumar Ray, [2]Tusar Dash
[1,2]Dept. of ENTC, OEC, BPUT, Bhubaneswar, Odisha, India

Abstract
With the increase of the Internet, the unauthorized copying and distribution of digital media is very difficult to stop, because the content can be played back in the original form and can be re-recorded and freely distributed. A solution for this problem is marking the media with a secret, imperceptible watermark. The media player can detect this mark and enforce a corresponding e-commerce policy. Audio watermarking is used as a multimedia copyright protection tool or as a system that embed metadata in audio signals. Here the challenge lies is to hide data representing some information without compromising the quality of the watermarked track. But to make sure that the embedded watermark is robust against removal attacks. Audio watermarking technique provides evidence for high audio quality.

Keywords
Audio Watermarking, Copyright Protection, Data Embeding

I. Introduction

Last decades have witnessed the wide spreading of the use and distribution of multimedia digital contents, mostly due to the diffusion of Internet connections with increasing bandwidth. In a world where the possibility to access information resides in the reliability of access to the World Wide Web, a need of finding an effective solution to the emerging need of defending the intellectual property (i.e. the so called copyright) from the hackers' attacks has arisen. All digital contents are concerned by this issue, from digital documents, to pictures and videos, as well as audio files (one of the first planetary case in this matter was indeed in the music field, with Napster and its service allowing users to download no matter what song in digital format for free and regardless to possible copyrights) [7]. From this background, the development of appropriate DSP techniques, like Digital Watermarking, has allowed the achievement (although not yet in a definitively acceptable way and with variable results among the different fields) of the purpose to control the distribution and protection of the various multimedia contents. Watermarking of digital data in such a controversial context has become an active and variegated research field, providing interesting results and solutions for safe data storage and transmission robust against malicious attacks.

The idea reported here is nevertheless not oriented to the improvement of data security but to an innovative use of audio watermarking, that is providing an assessment of the quality of an audio signal after an MP3 (MPEG-1 Layer III, ©) coding, transmission, and respective decoding. Although, as already reported, copyright protection has been the first application of Digital Watermarking, different uses of this technique have been recently proposed; digital fingerprinting and content dampers, encipherment of data (as a new stenography expression), or Audio Digital Watermarking applications for data retrieval are just few of the new watermarking utilizations. For what concerns the techniques used to protect the copyright avoiding unauthorised data duplications, the inserted watermark need to be detectable and not extractable without deterioration of the data themselves (i.e. it needs to be robust). In an opposite way, a fragile kind of watermark has been recently proposed by Professors Campisi, Carli, Giunta e Neri of Roma3 University of Rome to assess the Quality of Service (QoS) of a video communication in a blind way (which means being the original version of the watermarked file unavailable for extraction at the receiver side).

For what concerns the audio field, the main focus of the research on watermarking so far has been posed on either the direct marking of the audio stream or of its compressed version. While designing a watermarking scheme, the various characteristics of the file to be marked have to be kept into consideration as well as the particular application which the watermarking needs to realize. Many of the requirements for audio watermarking may be analogous to those of the picture watermarking, as for example transparency (i.e. the presence of the mark is supposed being not noticeable) and the robustness of the mark to manipulations and processing like compression, filtering, and A/D or D/A conversion. This means that in theory the mark should always be inserted in the data in a way which guarantees it is always not audible (a particularly strict requirement for music contents, as nearly perfect audio quality is generally considered as a fundamental point). In addition the mark embedding process should be chosen so to not blunt the mark robustness for what concerns eventual malicious attack expressly aimed to the mark removal. It is for those reasons that the audio watermarking technique proposed in the following for assessing the audio quality received after a coding-transmission process, is supposed to be applicable without the watermarking itself affecting neither the overall characteristics of the audio signal in which it is embedded nor the transmission effectiveness. In particular a fragile watermark has been inserted in an audio data stream audio of MP3-like type (MPEG-1 layer III) using a spread spectrum approach . On the receiving side, the watermark is extracted and compared to the original version (available on the decoding side as well, differently from the original version of the sent audio signal). The rationale supporting this approach is that the alterations suffered by the watermark as consequence of the coding and transmission process are likely to be the same as those suffered by the audio signal (or at least proportional to them) since both have been transmitted on the same communication channel. The QoS estimation is based on the calculation of the Mean Square Error (MSE) between the extracted watermark after the transmission and decoding and the original one before embedding, where as the non audibility of the mark is tested and assessed using the most recent perceptual techniques, like the PEAQ (Perceptual Evaluation of Audio Quality) that is the ITU (International Telecommunication Union) standard for measuring the objective quality of an audio file. In this case "perceptual quality" is obviously referred to the possibility of noticing the mark during the use of the music content and to the guarantee that the overall audio quality of the MP3 file is not corrupted after the watermark embedding. The proposed watermarking technique has been designed for application in multimedia communication systems. In fact, such a quality index can be used for a number of scopes in the newest telecommunications services ranging from the automatic and continuous feedback of the effective channel quality between user and radio mobile station to the real-time constant feedback to the server about the effective quality of service offered to the user.

II. Implementation

A. Watermark Embedding

At the source the audio watermarking procedure in question is depicted in fig. 1. It is possible noticing there how firstly the audio stream is divided in M frames. A 2 seconds long time slot has been chosen for each frame so that for audio tracks sampled at the standard frequency of 44.1 KHz, 88200 samples per frame are obtained. The watermark w[k] employed consists of a K bit long binary sequence. A pseudo-random noise vectors set PN (for each frames a distinct one is employed and all are known at the receiving side) is multiplied by the reference mark (unique for each transmission process and known at the receiving side as well).

$$w_i^{(s)}[k] = w[k] \cdot p_i[k], \quad i = 1, 2, \ldots, N$$

Here pi[k] is the i-th PN vector while wi(s)[k] is the spread version of the mark to be imbedded within the i-th block or frame. In an analogous way to what is generally known for other spread spectrum techniques, using multiple different spreading vectors PN ensures that the watermark embedding is performed differently from one frame of the host file to the other. It is worth noticing that in this way the watermark can be considered as non perceivable modification of the audio signal, being at the same time robust for what concerns the permanent bit errors caused by the physical structure of the network or by the management of the network itself (e.g. following to different paths in the transmission channel, multiuser interferences or overload factors and so on).

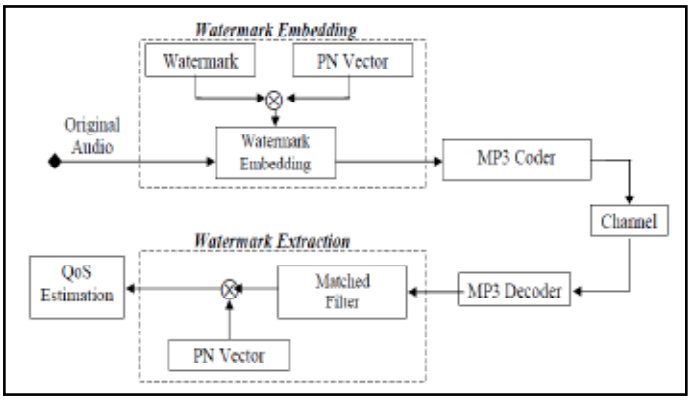

Fig. 1: Block Scheme of Tracing Watermarking for Channel Quality Assessment in Digital Audio Communications

B. Watermark Extraction

At the receiving side the decoding is performed together with the watermark detection (see once more fig. 1). For this purpose, after the decoded audio stream is filtered by a matched filter which extracts the watermark from the region of interest (the middlehigh frequency range) within the DCT transform of the received MP3 audio signal frames. The extracted watermark is then dispread so as to be comparable to the reference one. The matched filter is calibrated to the actual embedding procedure, so that it can detect the region where the spread watermark is located. In particular, the dispreading operation for the generic i-th block is performed following the relation:

$$\hat{w}_i[k] = \hat{w}_i^{(s)}[k] \cdot p_i[k]$$

It is thus possible to get an overall estimate of the received watermark wˆ[k] on a certain number M of transmitted frames starting from the following relationship

$$\hat{w}[k] = \frac{1}{M} \sum_{i=1}^{M} \hat{w}_i[k]$$

III. Results

With the aim of giving a concrete proof of the hypothesis formulated in the previous chapters, an extensive number of experimental tests have been carried out. The effective capacity of the mark in tracing the degradations of the audio signal -without affecting the audio quality by means of the embedding procedure- following a coding and transmission process had to be demonstrated.

During the experimentations, the mark has been embedded into the audio signal using the procedure described earlier. Considering the Nyquist-Shannon sampling theorem and the perceivable frequency range of the HAS, only audio files sampled at the frequency of 44.1 KHz e 48 KHz have been considered. A low payload file (for example a binary sequence as a bitmap logo) has been employed and embedded in every frame (as already said the frames are 2 seconds long each one) modulated by the parameter (the watermark robustness). As a trade-off between too large values (which would distort the audio signal while making the watermarking robust, i.e. useless for the scope proposed) and too short ones (making the watermark easily removable) a value 0.04 has been chosen.

The tests have considered the whole range of musical genres and compression rates, with preference given to the most common and used sampling rates which are 64, 96, 128, 160, and 192 kbps.

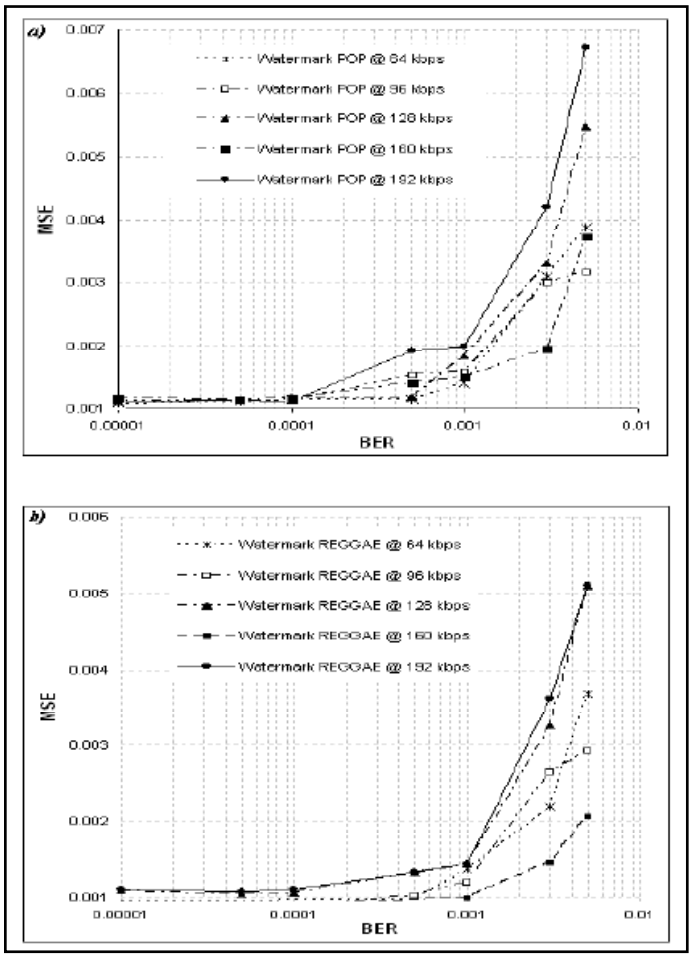

Fig. 2: MSE of the Watermark Extracted from MP3 Audio Signals at Different Compression Ratios

(a) POP music and (b) REGGAE music versus the BER of the channel for different compression ratios of the MP3 coder.

The following results on the other hand show that the presented digital fragile watermarking technique can be used to provide a reliable measure of the signal quality after an MP3 coding and transmission process, i.e. the starting hypothesis. Following the experimental procedure, the bit stream is sent on a noisy transmission channel simulated by means of a Poisson random

transmission error generator. This introduces a Bit Error Rate (BER) having value in the range from 10-5 to 5·10-3. In Fig. 2 and 3 the extracted watermark MSE with respect to the original one is considered, and it is related to the BER introduced within audio file of various musical genres.

The results from all those multiple genres (e.g. classical music, pop, reggae, swing) at multiple compression rates (from 64 to 192 Kbps) are reported. It is worth noticing that the MSE of the extracted watermark increases as the channel BER increases and as the bit rate decreases. This is coherent to the initial hypothesis that the MSE in the watermark reflects the degradation of the host audio signal, due to the increase of the errors introduced by the channel during transmission and to lower sampling rates. Moreover, as depicted by fig. 4 and 5, the corruption of the embedded watermark quality has the same behaviour (in terms of values trend) with respect to the audio quality degradation. As easily noticed by the graphs, the extracted watermark MSE is found to be strictly linked to the amount of error introduced within the audio bit stream: the increase in the BER value is followed by a proportional rising in the MSE of the watermark, which is in turn proportional to the rising in the MSE of the audio content.

IV. Conclusion

In this paper an application of audio digital signal processing techniques to digital watermarking is discussed. The audio watermarking technique has been proved to allow getting a blind assessment referred to the audio quality of musical signals received after being compressed in MP3 format, and transmitted on a normal, noisy channel. A fragile watermark is embedded into an audio uncompressed data stream by means of a spread spectrum approach. The watermark degradation can be therefore used to assess the decrease of audio quality to which the whole musical file has undergone. At the receiving side, the watermark is extracted and compared to the original counterpart. The QoS service assessment is performed using the Mean Square Error between the original reference watermark and the transmitted extracted one. Further developments of the idea are forecasted, by the joint consideration of multiple uses for the watermark.

References

[1] C. Wang, T. Chen, W. Chao,"A new audio watermarking based on modified discrete cosine transform of MPEG/ audio layer III", IEEE Int. Conf. on Networking, Sensing and Control, Vol. 2, pp. 984-989, 2004.

[2] F. Hartung, M. Kutter,"Multimedia Watermarking Techniques", Proc. of the IEEE, Vol. 87, No. 7, pp. 1079-1107, 1999.

[3] S. Cheng, H. Yu, Z. Xiong,"Enhanced spread spectrum watermarking of MPEG-2 AAC", IEEE Int. Conf. on Acoustics, Speech, and Signal Proc., ICASSP'02, Vol. 4, pp. 3728-3731, May 2002

[4] P. Bassia, I. Pitas, N. Nikolaidis,"Robust audio watermarking in the time domain", IEEE Trans. on Multimedia, Vol. 3, No. 2, pp. 232-241, June 2001

[5] M.D., Swanson, B. Zhu, A.H. Tewfik,"Current state of the art, challenges and future directions for audio watermarking", IEEE Int. Conf. on Multimedia Computing and Systems, Vol. 1, pp. 19-24 June 1999

[6] Y. Yaslan, B. Gunsel,"An integrated decoding framework for audio watermark extraction", Proc of the 17th Int. Conf. on Pattern Recognition, ICPR 2004, Vol. 2, pp. 879- 882, Aug. 2004

[7] I. Podilchuk, E. J. Delp,"Digital Watermarking: Algorithms and Applications", IEEE Sign. Proc. Mag., pp. 33-46, July 2001

[8] J. Haitsma, T. Kalker,"Speed-change resistant audio fingerprinting using autocorrelation", Proc. Int. Conf. on Acoustics, Speech, and Sign. Proc., Vol. 4, pp. 728-731.

Sanjay Kumar Ray received B.E degree in Electronics & Telecommunications Engineering from the Utkal University, Odisha, India in 1990 and M.Tech. degree in Electronics & Communications Engineering from Berhampur university in 2009 .He is currently working as Asst.Prof in OEC. His research interests include Wireless ad-hoc networks, Computer Communication& Netwcrking VLSI, Audio signal processing etc.

Tusar Dash received B.E. degree in Electronics & Telecommunications Engineering from the Utkal University, Odisha, India in 2003, and continuing the M.Tech. degree. He is currently working as a Lecturer in OEC. His research interests include Wireless ad-hoc networks, VLSI, Audio signal processing.