EE 683.01 Advanced Topics in Speech Processing Project Report
January 2001

AUDIO WATERMARKING USING DC LEVEL SHIFTING

Umut Uludag

M.Sc. Student
Electrical & Electronics Engineering Department
Bogazici University, Bebek, 80815
Istanbul, Turkey
umutuludag@ieee.org

Dr. Levent M. Arslan

Electrical & Electronics Engineering Department
Bogazici University, Bebek, 80815
Istanbul, Turkey
arslanle@boun.edu.tr

1. INTRODUCTION

Multimedia data are mainly in digital form due to advantages of digital domain compared to analog domain. These advantages can be grouped into several categories. Firstly, digital multimedia data authoring is easier and more robust to authoring errors. Editing and modifying data can be accomplished relatively easily. Secondly, digital multimedia data can be delivered over computer networks with almost no error and no interference. Thirdly, software instead of hardware processing of digital data increases the reconfigurability of systems ([1], [2]).

On the other hand, there are several disadvantages of digital media distribution. Digital data can be copied without any loss in quality and content. This poses a big problem for the protection of Intellectual Property (IP) rights of the copyright owners.

Watermarking is a solution to the previously defined problem. It can be defined as embedding digital data (generally information about origin, destination, access level of the data) imperceptibly into the host multimedia data. Several watermarking methods exist for different media like image, video, audio, text, polygonal models, etc. Watermark embedding domain can be spatial, frequency or both. In every method, embedding watermark data as robust as possible without being perceptible is desired. This requirement necessitates the utilisation of human sensory models (HVS: Human Visual System, HAS: Human Audible System…) and characteristics in watermark embedding process. More information about watermarking in general and audio watermarking can be found at [3], [4], [5], [6], [7] and [8].

2. TERMINOLOGY

Several techniques related to digital watermarking are:

Steganography relies on embedding the secret information in unsuspected data and generally used in secret point-to-point communication between communicating parties.
This technique is not robust to attacks.

Watermarking has the additional feature of robustness against attacks. Also, as opposed to steganographic techniques, destroying embedded data is difficult, even if the existence and method of embedding of digital data is known.

Data hiding and data embedding can be classified as techniques between steganography and watermarking.

Fingerprinting and labeling are watermarking techniques where watermark data is information about creator and recipient of digital data.

Bit-stream watermarking denotes watermarking of compressed domain data like video. Visible watermarks are visible patterns applied to images or video, mainly for Internet applications.

3. REQUIREMENTS

General watermarking requirements which are common to all digital media are as follows:

1) Watermark signal must carry as much information as possible, without being perceptible.

2) Watermark must be secure, i.e. only authorized parties must be able to access the watermark data. This security can be achieved by using cryptographic keys in watermark embedding and extraction.

3) Watermark must be robust against intentional and unintentional attacks. Compressing-decompressing, format change are examples to former kind of attacks. The latter kind are characterized by the intention of attacker to remove the embedded watermark from the host signal. For satisfying this requirement, a tradeoff analysis must be carried out between imperceptibility and robustness.

The requirements which are more application-dependant are:

1) Watermark decoding may or may not need the original host data. In general, watermark is more robust to attacks if the original host data is available at the receiver. Also, more data can be embedded as watermark in this case. But access to original data at receiver may be restricted in data monitoring applications, or due to large memory requirements, using original data for watermark recovery can be infeasible in applications like video watermarking.

2) Watermark embedding may need to work in real-time, especially for video watermarking.

3) Watermark signal may need to carry arbitrary information. Or it can carry information derived from a pre-specified dictionary. The former case requires the extraction of embedded arbitrary watermark; whereas in the latter case techniques like hypothesis testing can be applied to detect the presence of any watermark from the cited dictionary.

4. AUDIO WATERMARKING USING DC LEVEL SHIFTING

The audio watermarking method is explained below in three modules named watermark encoding, watermark decoding and watermark comparison.

4.1. Watermark Encoding

The method is based on shifting the DC levels of frames of input audio signal to positive or negative levels to indicate watermark bits 1 or 0, respectively. A secret key with 32 bits generate the binary watermark sequence associated with the copyright owner with that key. Linear Feedback Shift Registers (LFSR) are used in this process. Tap weights are selected to generate a binary sequence with maximum period (m sequence). Figure 1 shows such a binary watermark sequence with 64 bit length.

An input sample audio waveform is shown in Figure 2. (16 kHz sampling rate, 3.6 sec.)

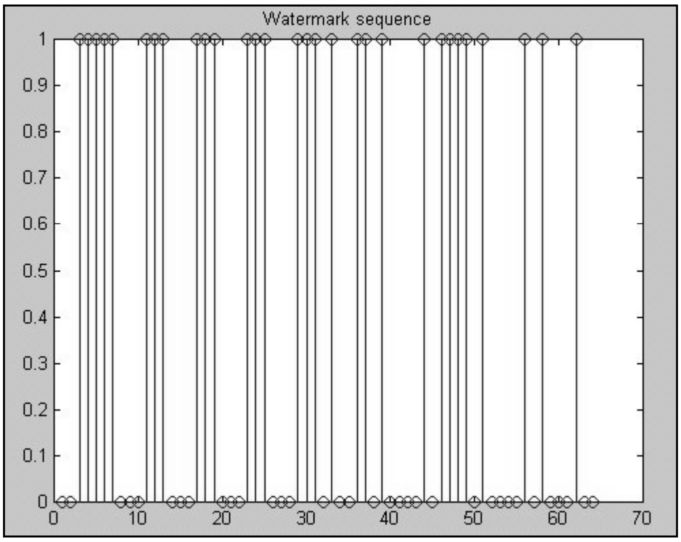

Figure 1. Original watermark sequence.

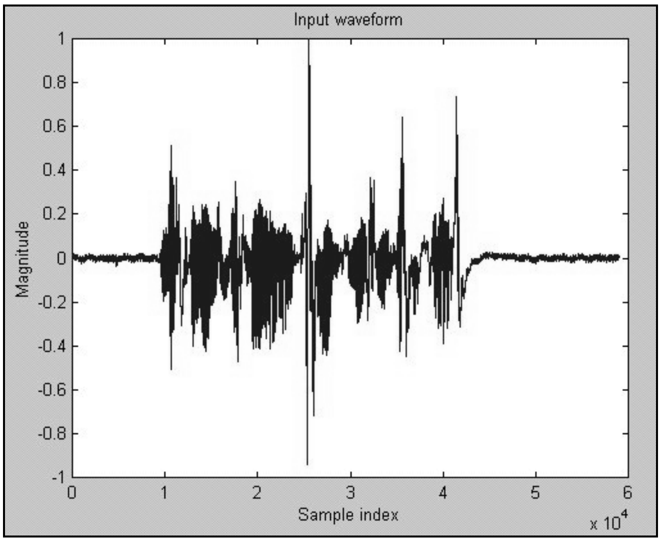

Figure 2. Input audio waveform.

Then, for every frame with duration 25 ms. in this audio file, frame means and frame powers are calculated. Frame means are shown in Figure 3 and frame powers are shown in Figure 4. Every individual frame is set to zero DC level by subtracting found frame means from audio samples in the associated frames.

This frame-wise DC zero audio sequence is processed to include watermark bits 0 and 1. Namely, the owner' s watermark sequence is embedded as follows: If the bit to embed is a zero, the corresponding frame' s DC level is shifted to a negative level with the value

$$level_0 = -DCBiasMultiplier * FramePower$$

If the bit to embed is a one, the corresponding frame' s DC level is shifted to a positive level with the value

$$level_1 = +DCBiasMultiplier * FramePower$$

The level shifting is made proportional to the power of the frame for inaudibility purposes. Figure 5 shows the watermarked waveform processed in the previously explained way.

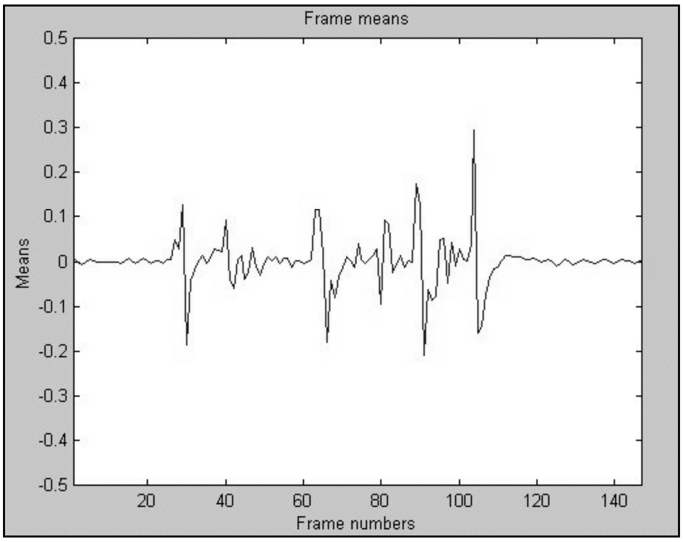

Figure 3. Frame means.

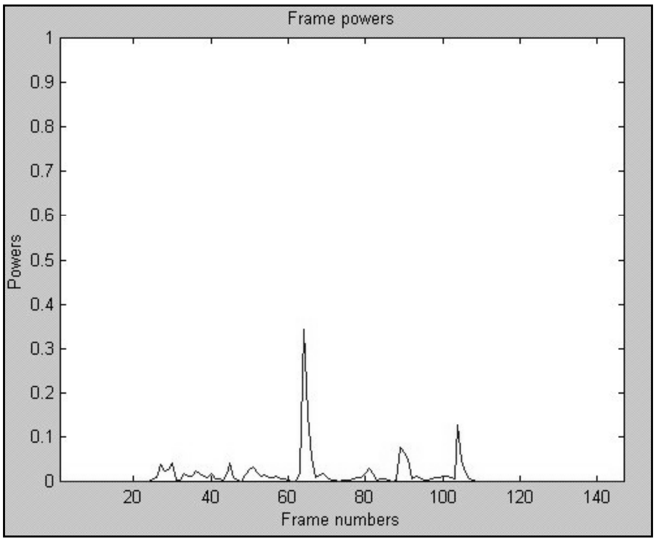

Figure 4. Frame powers.

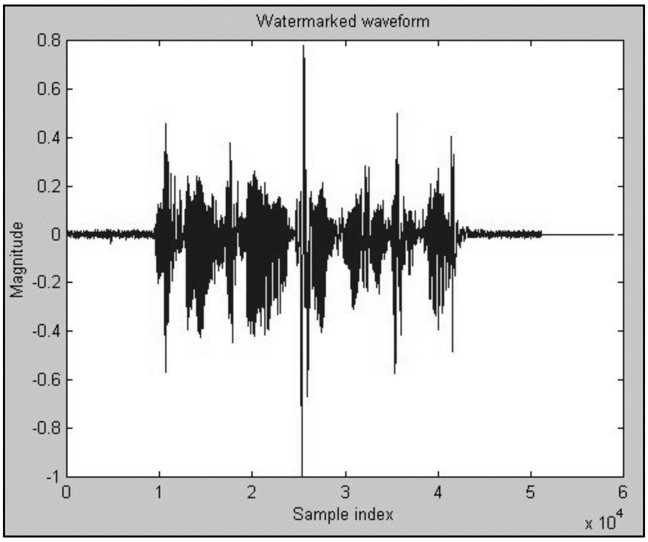

Figure 5. Watermarked waveform.

4.2. Watermark Decoding

In watermark decoding, the frame means of every frame in the input watermarked audio sample is calculated. Same watermark sequence is generally written onto the input audio sample multiple times (epochs). Epoch number increases if the duration of input audio is increased. For each of these epochs, the binary watermark sequence is decoded according to the sign of the frame means. Namely, if a frame has positive frame mean, the associated bit is 1; if a frame has a negative frame mean, the associated bit is 0.

The found frame means for the previosly watermarked audio file is shown in Figure 6.

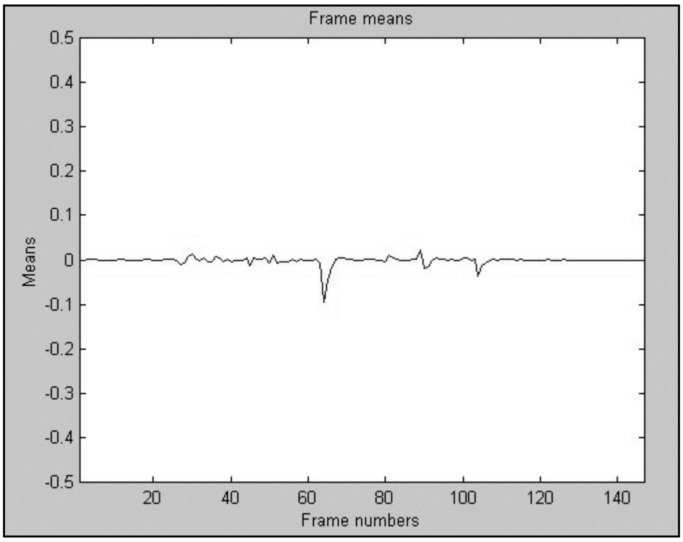

Figure 6. Decoded frame means.

The decoded watermark sequences are shown in Figure 7 and Figure 8 for epoch 1 and epoch 2, respectively.

4.3. Watermark Comparison

The original watermark sequence and the decoded sequence(s) are compared in this module. In the current implementation, if in any of the epochs, the Hamming distance between the sequences is less than 5 % of the binary sequence length, the watermark presence is declared. Otherwise, either the audio sample is not watermarked or it contains the watermark of another owner.

In the example audio given, these distances are 4 (6.25 %) and 1 (1.57 %) for two epochs. Therefore the encoded watermark is decoded accurately from the watermarked audio, without using original, unwatermarked audio.

5. CONCLUSIONS

The encoded watermark data is decoded correctly from the watermarked audio, without using the unwatermarked audio. This fact increases the applicability of the method to real world problems where the original, unwatermarked audio should only be available to the copyright owner. The audibility of the watermark signal itself is low. The source codes of the programs for previously defined modules are submitted with this report. A "README" file is also included.

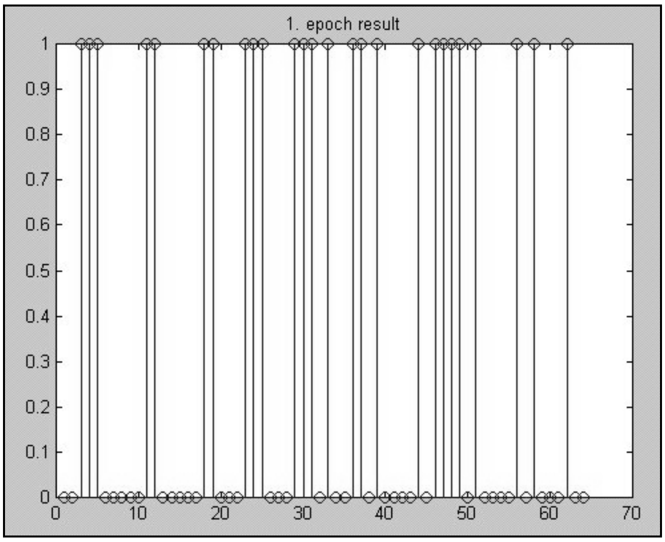

Figure 7. Decoded watermark sequence for 1. epoch.

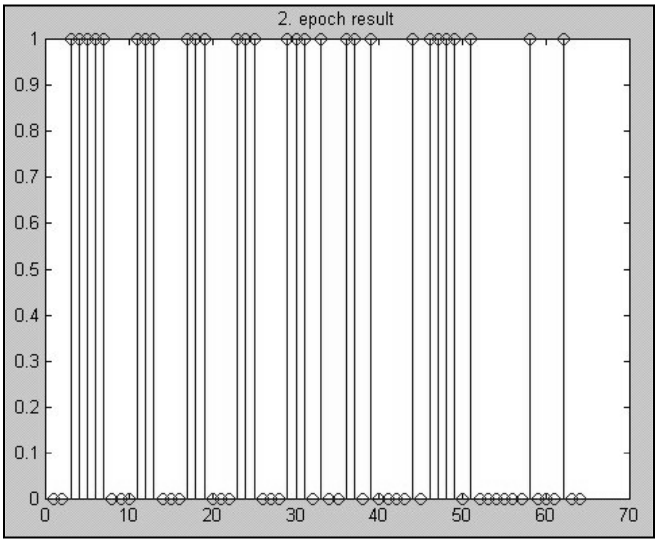

Figure 8. Decoded watermark sequence for 2. epoch.

REFERENCES

[1] F. Hartung and M. Kutter, "Multimedia watermarking techniques," *Proceedings of the IEEE,* vol. 87, no. 7, July 1999, pp. 1079-1107.

[2] M.D. Swanson, M. Kobayashi and A.H. Tewfik, "Multimedia data-embedding and watermarking technologies," *Proceedings of the IEEE*, vol. 86, no. 6, June 1998, pp. 1064-1087.

[3] M. Cooperman and S.A. Moskowitz, "Steganographic method and device," *U.S. Patent* 5 613 004, Mar. 18, 1997.

[4] F.A.P. Petitcolas, "Watermarking schemes evaluation," *IEEE Signal Processing Magazine*, vol. 17, no. 5, Sept. 2000, pp. 58-64.

[5] B.L. Greenberg, "Method and apparatus for the processing of encoded data in conjunction with an audio bradcast," *U.S. Patent* 5 379 345, Jan. 3, 1995.

[6] C.U. Lee, K. Moallemi, J. Hinderling, "Post-compression hidden data transport," *U.S. Patent* 5 687 191, Nov. 11, 1997.

[7] R.D. Preuss et al., "Embedded signalling," *U.S. Patent* 5 319 735, Jun. 7, 1994.

[8] M.D. Swanson, B. Zhu, A.H. Tewfik and L. Boney, "Robust audio watermarking using perceptual masking," *Signal Processing* 66(1998), pp. 337-355.