Visual Domain Audio Watermarking (VDAW) and Spectral Barcode Audio Watermarking (SBAW)

Alex Radzishevsky[1]

www.audiowatermarking.info

Original publication date: February 9th, 2014 (last edited: October 3rd, 2017)

ABSTRACT — This paper proposes a unusual, novel audio watermarking approach. The new technique combines audio signal processing with typical Optical Character Recognition (OCR) technique to create robust and efficient watermarking technology with high data rate.

I. INTRODUCTION

Watermarking, a form of steganography, is a process of hiding digital information in a carrier signal for the purpose of identification, annotation, signing and copyright. Various digital audio watermarking techniques and solutions have been developed. Commercial companies and scientific researchers active in this field have been proposing various digital signal-processing approaches and methods of hiding digital data within audio content in a inaudible way. Majority of the solutions developed so far provide means for secret watermarking in the sense that the result of the watermarking cannot be revealed by simple audio or visual inspection of the audio material. Most of such solutions represent commercial products and use proprietary algorithms.

One example of a sophisticated watermarking solution utilizing proprietary patented DSP approach is described in [1] and implemented in a set of software tools offered at www.audiowatermarking.info.

At the same time, there are different applications and use-cases, which don't require this high level of secrecy, but still require watermarking functionality, i.e. an ability of insertion and extraction of additional digital data (digital codes) into/from audio content. So far, there were not many such solutions available to general public, and those available were too simplistic and/or too obvious (such as mixing of voiced data into the audio content, or adding audible tonal signals carrying digital data by representing its individual bits by different durations of "beeps").

In this paper, a new method of digital data hiding inside acoustic content is proposed. The method is called Visual Domain Audio Watermarking (VDAW). Search in open informational sources did not reveal any prior references to the approach in this paper.

II. VISUAL DOMAIN AUDIO WATERMARKING AND SPECTRAL BARCODE AUDIO WATERMARKING

The idea of Visual Domain Audio Watermarking (VDAW) approach consists in embedding of graphically (visually) represented digital data (such as a barcode) within a visual representation of sound wave to allow further extraction (reading) of the embedded data from the visually represented sound using standard visual scanning tools (e.g. standard barcode scanner).

Spectrogram is a one particular example of different graphical sound wave representation forms available. Spectrogram (or sonogram) is a visually represented spectrum of frequencies in a sound as they vary with time. To be more precise, spectrogram image (refer to Figure 1) is a color/intensity map representing sequence of signal spectrum figures obtained using FFT decomposition of overlapping signal frames. It is a common way of graphical representation of audio content, which provides very high level of detailing and is easily readable visually. Additionally, spectrogram representation is very robust.

[1] Alex Radzishevsky is an inventor and developer of digital signal processing algorithms (www.audiowatermarking.info), and holds B.Sc. in Mathematics and Computer Science from the Technion (Israel Institute of Technology)

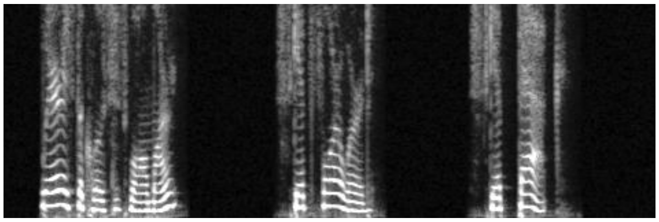

Figure 1. Exemplary spectrogram of human speech

So, why not "draw" a barcode on the sound spectrogram? This is exactly the idea behind Spectral Barcode Audio Watermarking (SBAW), one particular embodiment of VDAW. SBAW approach consists in embedding of graphical barcode (such as QR code, DataMatrix, or other 1D or 2D barcode) inside audio spectrogram. The idea of "drawing pictures" inside the sound spectrum was known for many years, but so far, it has never been used to carry barcodes with arbitrary digital data.

SBAW encoding process consists of the following three main stages:

1. transforming the audio signal into time-frequency domain (spectrogram),

2. "engraving" the barcode (carrying digital watermark payload) within the spectrogram by zeroing its particular regions corresponding to zero modules of the barcode,

3. backward synthesis of the modified audio spectrum into the output (encoded, watermarked) audio.

The encoding process is schematically depicted in Figure 2.

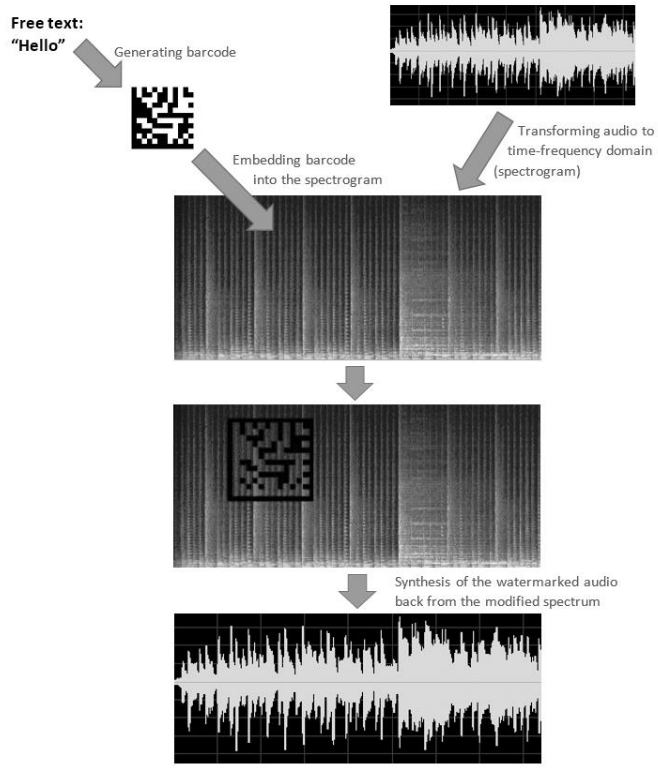

Figure 2. Watermarking process using the Spectral Barcode Audio Watermarking (SBAW) technique

Extraction of watermark from the audio signal comprises in:

- representing signal in a form of spectrogram (time-frequency domain),
- locating the barcode in the sound spectrogram (manually/visually or using automatic means),
- scanning the spectrogram image region containing the barcode with barcode scanner (either software or hardware, manually by a user or automatically using a special barcode detection and scanning tool),
- extracting the barcode data from the detected barcode.

The decoding process is schematically depicted in Figure 3.

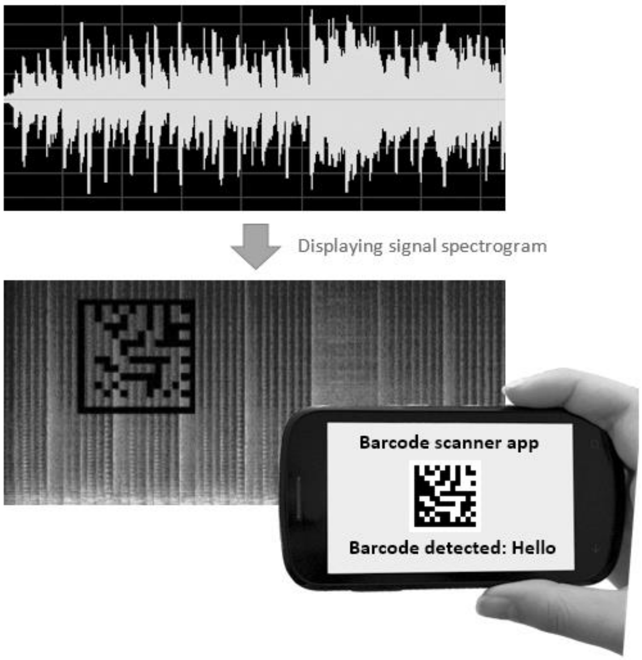

Figure 3. Watermark extraction from signal spectrogram using barcode reader

III. IMPLEMENTATION AND CONCLUSIONS

The proposed SBAW approach:

- is relatively easy to implement,
- provides robust watermarking, able to withstand even multiple lossy sound transformations such as MP3 encoding/decoding, time-stretching and pitch-shifting,
- is very scalable (depending on application and requirements, the barcode can be placed at different frequency regions and can last for different durations of time),
- provides very high watermarking data rate (especially with 2D barcodes),
- allows for a very time-accurate detection of the barcode location in the signal,
- provides nearly "inaudible" watermarking (especially if only high sound frequencies are used to carry the barcode).

In the simplistic implementation, SBAW method provides open ("not secret") watermarking in the sense that anyone can find the barcode using visual analysis of the sound spectrogram. If needed, secrecy can be achieved by obfuscation of the barcode data, obfuscation of the barcode picture, scrambling of spectrogram regions used as a watermark carrier and using other hiding methods. Additionally, various methods can be used to improve the robustness of the embedded data (such as adding pre-generated comfort noise in the carrier frequency region, etc.).

Different 1D and 2D barcodes can be utilized. The preferable barcode type is 2D, such as QR Code or Data Matrix [2]. Data Matrix barcode is likely the most preferable one as it is geometrically compact, provides high watermarking data rate, incorporates reliable error correction mechanism, is scalable and is convenient to embed into rectangular spectrogram picture.

SBAW-watermarked audio signal can undergo sound transformations such as MP3 encoding and decoding, but the barcode will still be detectable. At the same time, the audibility of the watermark remains minor when proper encoding settings are used.

The easiest manual way of barcode detection is scanning the barcode with barcode scanning application running on a smartphone. It is enough to point the smartphone camera on the computer screen showing the signal spectrogram in a proper geometrical proportion (refer to Figure 3), and the barcode will be detected successfully. This feature of the proposed technique is claimed to be innovative as it allows for watermark extraction from a specific signal representation (spectrogram) using no any proprietary tools, but only with standard barcode reader available for public.

Many alterations of the described basic idea can be applied to improve detectability, watermarking data rate, inaudibility (transparency) and robustness of the watermark.

SBAW is only one embodiment of many other possible embodiments of much more general VDAW approach. The very same basic idea is potentially applicable to other kinds of "visual domains" (visual sound representations) having forward and backward single-valued transforms.

Although simplicity, effectiveness and novelty of the proposed approach can be a good sign of invention, its patentability remains questionable in the light of the prior art known in the field of barcodes, visual representation of audio signals and modulation techniques. For the above reason, and in order to contribute to public knowledge, the author decided to openly publish description of the proposed approach within the present article and make it publicly available.

REFERENCES

[1] A. Radzishevsky, Watermark embedding and extraction, U.S. Patent 8,116,514, February 14, 2012.

[2] Data Matrix barcode, https://en.wikipedia.org/wiki/Data_Matrix