A Multiple Watermarking Scheme for Digital Audio Signals

Maria Chroni
mchroni@uoi.gr
Department of Computer Science & Engineering
University of Ioannina
Ioannina, Greece

Stavros D. Nikolopoulos
stavros@cs.uoi.gr
Department of Computer Science & Engineering
University of Ioannina
Ioannina, Greece

Iosif Polenakis
ipolenak@cs.uoi.gr
Department of Computer Science & Engineering
University of Ioannina
Ioannina, Greece

Vasileios Vouronikos
v.vouronikos@uoi.gr
Department of Computer Science & Engineering
University of Ioannina
Ioannina, Greece

ABSTRACT

Watermarking is a method for the verification of the authenticity of a digital object. In this work we investigate the watermarking of audio signals utilizing watermarked images. We present a multiple watermarking scheme for the watermarking of digital audio signals. The proposed technique exploits an image as a watermark which has been already watermarked with a self inverting permutation. The watermarked image, is then embedded in the LSB values of the digital audio signal. By conducting a series of evaluation experiments, including an attack vector consisting of cropping, noise, and re-sampling attacks, we investigate the robustness of the proposed multiple watermarking technique and prove its potentials against various factors affecting each type of attack.

CCS CONCEPTS

• **Mathematics of computing** → *Coding theory*; • **Security and privacy** → **Information-theoretic techniques**; **Authentication**; **Digital rights management**; **Economics of security and privacy**; **Social aspects of security and privacy**; • **Applied computing** → *Investigation techniques*; **Data recovery**; **Law**.

KEYWORDS

algorithms, security, information hiding, watermarking

ACM Reference Format:
Maria Chroni, Stavros D. Nikolopoulos, Iosif Polenakis, and Vasileios Vouronikos. 2022. A Multiple Watermarking Scheme for Digital Audio Signals. In *26th Pan-Hellenic Conference on Informatics (PCI 2022), November 25–27, 2022, Athens, Greece.* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3575879.3576005

PCI 2022, November 25–27, 2022, Athens, Greece

ACM ISBN 978-1-4503-9854-1/22/11…$15.00
https://doi.org/10.1145/3575879.3576005

1 INTRODUCTION

The term digital watermarking expands to a multitude of digital objects including, among others, digital images, software, video files, audio files, etc. The purpose of watermarking is to cover a piece of information, but in such a way that it can be easily retrieved if someone wants to retrieve it. However, the information that has been hidden into a digital object should not noticeably affect its quality, but should somehow become an integral part of it. Digital object watermarking has its basis on the elementary procedures of embedding and extracting an information into the digital object under consideration, where both these procedures are required to be as efficient, simple, and reliable as possible. In particular, considering the watermarking procedure of a digital object (in our case an audio signal S), an audio watermarking technique seeks to introduce modifications that are primarily undetectable by the human ear into the audio's data in order to incorporate a distinctive identifier, i.e., the watermark w [7]. Moreover, the main target of digital object watermarking remains that whether a digital object has been watermarked with a watermark, let w, and the digital object is transferred over the Internet, then w is also transferred with the copy of the digital object ensuring the maintenance of copyright protection [7].

1.1 Audio Watermarking

Regarding the watermarking procedure of a digital audio signal, as described in [7], given a digital object O (in our case an audio signal S) into which we want to embed a watermark w there should be developed an embedment procedure, let $Embed(O, w)$, which implements the embedding of watermark w into the digital object O resulting thus to the watermarked digital object O_w (in our case, considering the audio signal S, the S_w). However, it is notable to refer that once a digital object has been watermarked it is required for the watermark to be able to be preserved to the greatest extent possible, despite any alterations that the watermarked object may undergo. Similarly, for the reverse procedure, i.e., the one of extracting the watermark w from a watermarked digital object O_w, there should be developed the corresponding extraction procedure, let, $Extract(O_w)$, which, given the watermarked digital object O_w extracts, by implementing the reverse procedure, the watermark w. To this point, it is also notable to refer that the main concern regarding the watermarking of digital objects, considering in our case the

digital audio signals, audio watermarking requires the consistent and precise location and extraction of the embedded watermark w from the watermarked audio signal S_w even in the case where S_w has undergone a set of attacks, e.g., cropping, noise addition, re-sampling, etc.

1.2 Related Work

Audio watermarking techniques have been classified into time domain (or spatial domain), frequency domain, and hybrid domain. The time domain techniques mainly utilize Least Significant Bit (LSB) substitution and echo-hiding techniques. In LSB technique, the audio signal is sampled at 8 or 16 kHz and divided into frames, and the LSB of each frame is replaced by the watermark bit [4]. The robustness depends on the number of bits that are being replaced in the host signal. To increase the robustness and imperceptibility, various modifications in LSB technique have been proposed by changing the embedding positions [9, 11]. The audio signal has real values as samples, if converted to an integer will degrade the quality of the signal to a great extent. The time-domain watermarking techniques are found in [3, 5, 10, 12, 15, 16]. Generally, time-domain watermarking techniques are simple and less complex but suffers with low robustness [13]. In frequency domain audio watermarking techniques the watermark is embedded in frequency domain with any transformation. This technique is a bit complex to implement and it requires more computation in comparison with the spatial domain, because in frequency domain the robustness is much better. There are various techniques for embedding the watermark in transform domain for digital audio signal such as FFT, DCT, DWT and SVD [2]. Several methods have been proposed in the literature for embedding an image in an audio. Authors of [14] propose a method of digital watermarking using low frequency components, pseudo numbers and a binary image file, whereas authors of [6] use a digital image as watermark and present a method of embedding the watermark image in the LSB plane of the audio file. In [1] authors propose an an SVD audio watermarking approach using chaotic encrypted images, where recently authors of [17], proved that the proposed watermarking algorithm by Al-Nuaimy et al. falls to two ambiguity attacks where the extracted watermark is not the embedded one but determined by the reference watermark.

1.3 Our Approach

In this work we propose a multiple watermarking approach for the watermarking of digital audio signals. In our approach we are based upon the “image-to-audio" watermarking technique utilized in many approaches during the recent years through literature and augment this approach utilizing an already watermarked image. In the proposed technique we leverage the duality of the hidden information considering both the encoding a digital image (which is already watermarked) as well as the watermark itself, encoding the watermarked image in the LSB of the values of a digital audio signal. The utilization of the LSB provides the ability of tamper detection and the effectiveness as also the efficacy of the overall approach regarding the embed and the extract procedures required respectively for the embed and the extract of a watermark into a digital audio signal.

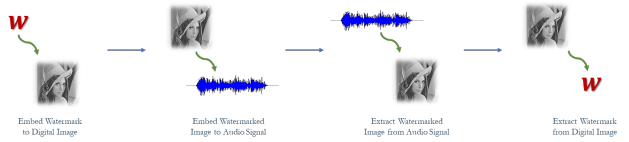

Figure 1: Overview of the procedures deployed for the embed end the extract of multiple watermarks into and from digital audio signals.

2 THE MODEL

In this section we describe the proposed algorithms for audio watermarking and watermark extraction to and from an audio signal. The proposed multiple watermarking scheme for digital audio signals incorporates two main procedures, namely the embed and the extract of information, in our case watermarked images, into/from, digital audio signals. The structure of these two procedures contains the embed of a watermark into a digital image and the consequent embed of the watermarked image into a digital audio signal, while the reverse procedure contains the extraction of the watermarked image from the digital audio signal that it has been embedded into in order to extract from it the embedded watermark. In Figure1 it is presented an illustrative example of how these procedures are deployed in order to implement our proposed multiple watermarking scheme for audio signals.

2.1 Embedding a Watermark into a Digital Image

The procedure that implements the embedment of a watermark into a digital image incorporates the watermarking procedure utilizing the properties of Self-Inverting Permutations presented in [8]. In particular, we deploy the 2DM representation of Self-Inverting Permutation (SiP), which encodes an integer w and is used to locate specific cells of an image I in order to embed the SiP that represents the watermark w into the image I. In particular, we select an integer w and utilize the 2DM representations of the SiP that encodes it, to locate the corresponding cells of the image. In Algorithm 1 we describe the steps followed for the embedment of the watermark w into a image I producing its corresponding watermarked image I_w.

2.2 Embedding a Watermarked Image into an Audio Signal

The procedure that implements the embedment of a watermarked digital image I_w into a digital audio signal S incorporates the utilization of the Least Significant Bit (LSB) modifying the technique proposed in [6] and in the approach proposed in [18] in order to embed the information that encodes the watermarked image I_w into a digital audio signal S constructing thus the watermarked audio signal S_w. In particular, we utilize the values of the cells of a matrix that represents the watermarked gray-scale image I_w of size $m \times n$ and transform it into an $1D$ matrix of size $1 \times \ell$, where $\ell = m \times n$. To this point we should note that since I_w is a gray-scale image, it holds that the values of its corresponding matrix are in the range $[0, 255]$. This matrix then is converted to a matrix of size $\ell \times 8$ where each of its rows contains 8 cells that represent the

Data: Image I, Integer Number w to be Embedded in the Image I
Result: Watermarked Image I_w
Construct 2DM representations of $Sip(w)$;
for *all cells in the 2DM representations* **do**
  Compute the FFT on each cell;
  Extract the corresponding *magnitude* M_{ij} and *Phase* P_{ij} matrices ;
  Place the Red-Blue Annuli on M_{ij} ;
  Mark the Red-Annuli of M_{ij};
  Combine M_{ij} with P_{ij};
  Compute the iFFT on $C'_{i,j}$;
end

Algorithm 1: Embed_Watermark_to_Image Algorithm.

Data: Audio Signal S, Watermarked Image I_w
Result: Watermarked Audio Signal S_w
$S_w \leftarrow S$;
Let m, n the size of the image I_w;
Let $M[i, j]$ the cell (i, j) in the $I_w : 0 \leq i \leq m$ and $0 \leq j \leq n$;
Construct matrix A of size $1 \times \ell : \ell = m \times n$;
$counter \leftarrow 0$;
for $i = 1 : 1 : m$ **do**
  for $j = 1 : 1 : n$ **do**
    $A[counter++] \leftarrow M[i, j]$
  end
end
Construct matrix B of size $\ell \times 8$ by containing the cells of A by their corresponding binary (8−bit) representation;
Construct sequence σ of length $\kappa = \ell \times 8$ that contains row-by-row the elements of B;
Select κ samples $s_1, s_2, \ldots, s_k \in S_w : k = |\sigma| \leq |S_w|$
for $i = 1 : \frac{|S_w|}{\kappa} : |S_w|$ **do**
  $LSB(s_i \in S_w) \leftarrow \sigma_i$;
end

Algorithm 2: Embed_Image_to_Audio Algorithm.

Data: Watermarked Audio Signal S_w, Integer Numbers m, n
Result: Watermarked Image I_w of size m, n
Construct an empty sequence σ of length $m \times n \times 8$;
Select κ samples $s_1, s_2, \ldots, s_k \in S_w : k = |\sigma| \leq |S_w|$
for $i = 1 : \frac{|S_w|}{\kappa} : |S_w|$ **do**
  $\sigma_i \leftarrow LSB(s_i \in S_w)$;
end
Construct matrix B of size $\ell \times 8 : \ell = m \times n$;
Parse sequence σ and place each 8 elements (binary digits) into each row of matrix B ;
Construct matrix A of size $1 \times \ell : \ell = m \times n$ containing the integer value of each cell of matrix B (8−bit binary representation) ;
Construct matrix M of size $m \times n$;
$counter \leftarrow 0$;
for $i = 1 : 1 : m$ **do**
  for $j = 1 : 1 : n$ **do**
    $M[i, j] \leftarrow A[counter++]$
  end
end

Algorithm 3: Extract_Image_from_Audio Algorithm.

Data: Watermarked Image I_w,
Result: The Integer w that is embedded in I_w
Construct 2DM representation of $Sip(w)$;
for *all cells in the 2DM representation* **do**
  Compute the FFT on each cell;
  Extract the corresponding *magnitude* M_{ij} and *Phase* P_{ij} matrices ;
  Place the Red-Blue Annuli on M_{ij} ;
  For this line select $C'_{i,j} : Avg_{B_{ij}} - Avg_{A_{ij}} = min\{Avg_{B_{ij}} - Avg_{A_{ij}}\}_{j=1}^{j=\ell(\pi^*)}$;
  Construct the SiP;
  Decode the embedded integer w;
end

Algorithm 4: Extract_Watermark_from_Image Algorithm.

$8 - bit$ that correspond to the binary representation of the value on the cell of each row of the $1D$ matrix, note that 8 bits are adequate to represent the values in the range $[0, 255]$. The produced matrix contains in its first column the Most Significant Bit of the represented number while in its last column it contains the Least Significant Bit. To this point, and having a matrix of $\ell \times 8$ elements, we construct a sequence, let σ of length $\ell \times 8$, where it contains sequentially the rows of the previous matrix, i.e., a sequence of 0 and 1. Respectively, with $step = \frac{|S|}{\kappa}$ we select $\kappa = \ell \times 8$ samples from the audio signal and insert into the LSB of each sample the corresponding digit of the sequence σ, in order to create finally the watermarked signal S_w. In Algorithm 2 we describe the steps followed for the procedure of embedding a watermarked image I_w to an audio signal S resulting to the watermarked audio signal S_w.

2.3 Extracting a Watermarked Image from a Watermarked Audio Signal

The procedure that implements the extract of a watermarked digital image I_w from a watermarked digital audio signal S_w incorporates again the utilization of the Least Significant Bit (LSB) modifying the technique proposed in [6] in order to extract the information that encodes the watermarked image I_w from a watermarked digital audio signal S_w. In particular, given a watermarked digital audio signal S_w and known the integers m, n that correspond to the information required regarding the extraction of the embedded watermarked image I_w, with $step = \frac{|S|}{\kappa}$ we select $\kappa = \ell \times 8$ samples from the audio signal, note that $\ell = m \times n$, and extract the LSB from each sample, in order to construct a sequence σ of values 0 and 1. Considering blocks of 8 values in the sequence we construct a matrix of size $(\frac{|\sigma|}{8}) \times 8$ in order to store the values of the sequence. Then, following the reverse procedure, since each row of the matrix represents a number in the range $[0, 255]$, where using this binary representation, we construct a matrix of $\frac{|\sigma|}{8}$ rows and one column storing the decimal representations that correspond to the numbers represented by their binary form in each row (8-bit binary representation). This matrix then is transformed into a 2D matrix of size $m \times n$ selecting each 8 sequential values and setting them as a row in the matrix. The resulting matrix then contains the values of each cell of the watermarked image I_w. In Algorithm 3 we describe the steps followed for the procedure of extracting a watermarked image I_w from a watermarked digital audio signal S_w.

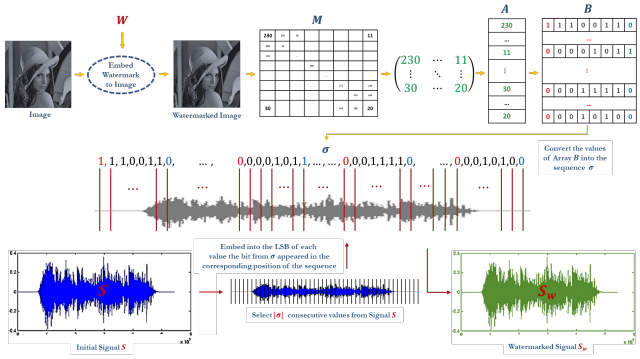

Figure 2: Embed Procedure Architecture

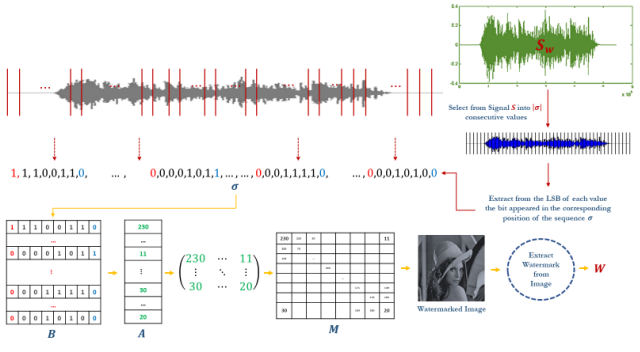

Figure 3: Extract Procedure Architecture

2.4 Extracting a Watermark from a Watermarked Digital Image

Similarly to the ebbed procedure presented, we deploy the 2DM representation of the SiP [8] that encodes the integer w that is embedded in the image I to locate the specific cells of image I that contain the watermark in order to extract them from the watermarked image I_w. In particular we incorporate again the 2DM representations of the SiP that encodes the integer w we previously embedded in the cells of I through the corresponding procedure, to locate the cells of the image and extract the embedded information. In Algorithm 4 we describe the steps followed for the procedure of extracting the watermark w from a watermarked digital image I_w.

2.5 System Architecture and Deployment

The proposed scheme for embedding and extracting information into and from digital audio signals integrates sequentially the embed and extract procedures, respectively. For the embedding of information into the digital audio signal, we embed a watermark w into a gray-scale digital image I producing the watermarked image I_w, and consecutively embed the watermarked image into the audio signal S producing the watermarked audio signal S_w. In Figure 2 it is illustrated the overview of the architecture of the embed procedure incorporating the sequential utilization of Algorithm 1 to embed the watermark w into the frequency domain of image I producing the watermarked image I_w, and Algorithm 2 to embed

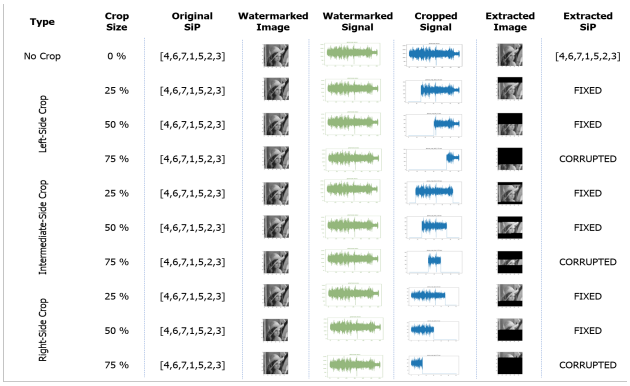

Figure 4: Crop attack demonstration.

the values of the matrix-representation of image I_w into the LSBs of particular values of the audio signals S in order to produce the watermarked audio signal S_w. For the extract of information from the digital audio signal there are utilized sequentially the procedure of extracting the watermarked image I_w from a watermarked audio signal S_w and consecutively the procedure of extracting the watermark w from the extracted watermarked image I_w. Similarly, in Figure 3 it is illustrated the overview of the architecture of the extract procedure incorporating the sequential utilization of Algorithm 3 to extract the values from the LSBs of the watermarked audio signal S_w that correspond to the matrix-representation of image I_w and produce the resulting watermarked image I_w, and the Algorithm 4, to extract the watermark w that has been embedded into the frequency domain of the image.

3 EVALUATION

In this section we provide an experimental evaluation of our proposed approach for the embedding and the extracting of information into and from digital audio signals, discussing our methodology, the experimental setup followed for the evaluation of our model, and the exhibited results that prove the potentials of our approach.

3.1 Attack Vector against Watermarked Digital Audio Signals

Some common processes that alter a digital object, and thus may alter the watermark, are signal conversion from digital to analog (and vice versa), sampling, quantization, compression and decompression, Gaussian noise, etc. Next, we discuss the types of attacks, namely, crop attack, noise attack, and down-sampling attack performed in order to evaluate the proposed multiple watermarking scheme for digital audio signals.

Through the crop attack we investigate the potentials of our proposed model for the embed of multiple watermarks into a digital audio signal regarding the case where the derived watermarked audio signal S_w has been cropped. We distinguish three types of crop, namely left-side crop, intermediate-side crop, and right-side crop where the watermarked audio signal has been cropped from its start until a specific part of it, a chunk of the signal has been removed from its middle, and the case where it has been cropped from a specific point until its end, respectively. In all these cases we

consider the size of the signal that has been cropped to correspond to an 25%, a 50%, and a 75% of the size of the initial watermarked audio signal S_w. In Figure 4 it is provided an illustrative example over each case of cropping attack and the corresponding size of cropped signal, where it is also depicted the corresponding watermarked image I_w extracted over each case.

Next, through the noise attack we investigate the potentials of our proposed model for the embed of multiple watermarks into a digital audio signal regarding the case where the derived watermarked audio signal S_w has been subjected to noise attacks. We used random noise based on Gaussian distribution with three types of variance, namely $\sigma^2 = 0.001$, $\sigma^2 = 0.01$, $\sigma^2 = 0.5$. In those attacks we increasingly add noise in the watermarked signal S_w in order to explore the effect on the watermarked image we embedded I_w. Noise attacks are a common filter attack that malicious users utilize in order to destroy fragile watermarks and claim ownership of the digital object in question.

Finally, through the down-sampling attack we investigated several down-sampling ratios to the watermarked audio signal S_w so as to explore the effect of down-sampling to the embedded watermarked image I_w. We investigate down-sampling ratios as they are more common to occur utilizing three types of down-sampling ratios SR', namely half-ratio ($SR' = SR/2$), quarter-ratio ($SR' = SR/4$) and eighth-ratio ($SR' = SR/8$), where SR and SR' correspond the initial and the down-sampling ratios of the watermarked audio signal S_w, respectively. Down-sampling attacks are a type of attacks where a malicious user re-samples an audio signal so as to reduce its size with the intention to get rid of embedded watermark values.

3.2 Methodology and Experimental Design

Throughout a series of experiments for various factors regarding each type of attack, we focus on the investigation of the cases where either the embedded watermarked image I_w, or the watermark embedded into it i.e., w, can be extracted and in particular for the case of the extracted watermark, if it has been damaged by the attack, whether it can be reconstructed utilizing the Self-inverting Permutation properties described in [8]. In the experimental design followed for the evaluation of our proposed model we attest its robustness in extracting the embedded watermarked image I_w and the watermark embedded to that image w when they are extracted from the watermarked signal S_w after the performance of several attack types regarding various factors that characterize each one.

In our experiments we utilized as watermark the integer number 5 that is encoded by the Self-inverting Permutation $\{4, 6, 7, 1, 5, 2, 3\}$ and embedded this SiP into the image of *Lena* of dimensions 128×128. Then, this image is embedded into a digital audio signal S, where in our case we distinguished three types of signals of length ranging from 30 seconds to 1 minute including only music, music and song, and voice only tracks. The types of attacks performed on the audio files were the crop attack, the noise attack, and the down-sampling attack. For the case of cropping there where investigated in each case of audio signal the crop of the track by its left side (left-side crop), its right side (right-side crop) and the intermediate-side (cropping from both left and right sides), for the case of noise attack the insertion of noise of various variances into the audio signal, and for the case of down-sampling the re-sampling of the audio signal using the half, the quarter and the eightieth of its initial sample rate. To this point, we should note that in all these attack cases our main target is to evaluate the effectiveness of our model regarding the successful extraction of the information embedded into the audio signal considering both I_w and w.

3.3 Discussion over the Exhibited results

The experimental results exhibited after the evaluation of our proposed model for audio signal watermarking utilizing multiple watermarks, i.e., the watermarked image I_w and the watermark w are presented in Table 1. The series of experiments are divided into three main categories regarding the underlying type of attack deployed. Utilizing three audio signals (tracks in .wav format), namely `music.wav`, `piano.wav`, and `crowd.wav`, including music and song, music only, and voice only, respectively, we performed the three types of attacks, namely *cropping*, *noise* and *down-sampling*, in order to evaluate our proposed technique.

For the case of crop attack the three tracks we cropped using the three types of crops (i.e., left-side, intermediate-side, and right-side crop, respectively), cropping the track by a magnitude of 25%, 50%, and 75%, respectively in each case. As we can observe from the results depicted in Table 1, for the cropping of the track by a magnitude of 25% and 50% (i.e., an 75% and a 50% of the track has been left available, respectively, for the watermark extraction procedure), the proposed technique achieves an 100% successful extraction of both the watermarked image I_w and the watermark w. However, for the case where the 75% of the track has been cropped (i.e., only an 25% of the has been left available for the watermark extraction procedure) in presence of a left-side crop only the watermark w embedded in the watermarked image I_w can be extracted after the attack, where for the cases of intermediate-side and right-side crop neither the watermarked image I_w nor the watermark w can be extracted successfully. Thus, through an overall evaluation of the proposed technique against the crop attacks, it exhibits an 66.6% successful extraction of both watermarked image I_w and watermark w in all types of crop attack for all the crop sizes (i.e., 25%, 50%, and 75%), where in particular for the case of left-size crop it is notable to refer that the proposed technique achieves an 100% successful extraction of the watermark w even in the extreme case where only an 25% of the track has been left available for the extraction procedure after the crop attack (i.e., 75% crop).

Moreover, for the case of noise attacks, as we can observe from the exhibited experimental results presented in Table 1 the proposed technique achieves an 100% successful extractions of both the watermarked image I_w and watermark w in all the cases. In particular, we investigate for all the three digital audio signals the deployment of noise attack utilizing the three types of variance, i.e., $\sigma^2 \in \{0.001, 0.01, 0.5\}$, where the PSNR value achieved an average of 65.09.

Finally, for the case of down-sampling attacks, as we can observe from the exhibited experimental results presented in Table 1, the proposed technique achieves also an 100% successful extractions of both the watermarked image I_w and watermark w in all the cases where the sample rate SR' has been divided to the half, the quarter, and the eightieth, i.e., $SR' = SR/2$, $SR' = SR/4$, and $SR' = SR/8$, respectively, where the PSNR value achieved an average of 63.09.

	Crop Attacks (Left, Intermediate, Right)				Noise Attacks			Down Sampling Attacks		
Audio	**Size**	I_w/w **(L)**	I_w/w **(I)**	I_w/w **(R)**	σ^2	**PSNR**	I_w/w	SR'	**PSNR**	I_w/w
music	25.00%	✓ / ✓	✓ / ✓	✓ / ✓	0.001	62.19	✓ / ✓	SR / 2	62.19	✓ / ✓
	50.00%	✓ / ✓	✓ / ✓	✓ / ✓	0.01	62.19	✓ / ✓	SR / 4	62.19	✓ / ✓
	75.00%	✗ / ✓	✗ / ✗	✗ / ✗	0.5	62.19	✓ / ✓	SR / 8	55.70	✓ / ✓
piano	25.00%	✓ / ✓	✓ / ✓	✓ / ✓	0.001	66.54	✓ / ✓	SR / 2	66.54	✓ / ✓
	50.00%	✓ / ✓	✓ / ✓	✓ / ✓	0.01	66.54	✓ / ✓	SR / 4	66.54	✓ / ✓
	75.00%	✗ / ✓	✗ / ✗	✗ / ✗	0.5	66.54	✓ / ✓	SR / 8	60.08	✓ / ✓
crowd	25.00%	✓ / ✓	✓ / ✓	✓ / ✓	0.001	66.54	✓ / ✓	SR / 2	66.54	✓ / ✓
	50.00%	✓ / ✓	✓ / ✓	✓ / ✓	0.01	66.54	✓ / ✓	SR / 4	66.54	✓ / ✓
	75.00%	✗ / ✓	✗ / ✗	✗ / ✗	0.5	66.54	✓ / ✓	SR / 8	60.08	✓ / ✓
avg	N/A	66.6% / 100%	66.6%	66.6%	N/A	65.09	100%	N/A	63.09	100%

Table 1: Experimental results after the deployment of Crop, Noise, and Down-sampling attacks utilizing the proposed technique.

4 CONCLUSION

In this work we presented a multiple watermarking scheme for the watermarking of digital audio signals utilizing watermarked images. The proposed technique leverages the duality of the hidden information by means of embedding into the LSB of the values of digital audio signal both a digital image (watermarked image) as also the watermark embedded in the digital image. We attested the robustness of the proposed multiple watermarking technique by a series of evaluation experiments including an attack vector consisting of cropping, noise, and re-sampling attacks investigating its potentials against various factors on each type of attack.

The promising results achieved through the evaluation of the multiple watermarking scheme for the watermarking of digital audio signals utilizing watermarked images proposed in this work prove to a further extent the potentials of the proposed model for digital audio watermarking leading us to insist for a further investigation of its insights and the explore of further properties that could improve its capabilities. Moreover, through our research scope for future work, we plan to investigate also its potentials when applied into audio signals utilizing already watermarked audio signals instead of digital images as also the modification of its architecture in order to develop watermarking schemes that could effectively embed watermarked audio signals to digital images.

ACKNOWLEDGMENTS

This research was supported by project “Dioni: Computing Infrastructure for Big-Data Processing and Analysis" (MIS No. 5047222) co-funded by European Union (ERDF) and Greece through Operational Program “Competitiveness, Entrepreneurship and Innovation", NSRF 2014-2020.

REFERENCES

[1] Waleed Al-Nuaimy, Mohsen AM El-Bendary, Amira Shafik, Farid Shawki, Atef E Abou-El-azm, Nawal A El-Fishawy, Said M Elhalafawy, Salaheldin M Diab, Bassiouny M Sallam, Fathi E Abd El-Samie, et al. 2011. An SVD audio watermarking approach using chaotic encrypted images. *Digital Signal Processing* 21, 6 (2011), 764–779.

[2] Jaya Bajpai and Arashdeep Kaur. 2016. A literature survey-Various audio watermarking techniques and their challenges. In *2016 6th International Conference-Cloud System and Big Data Engineering (Confluence)*. IEEE, 451–457.

[3] Paraskevi Bassia, Ioannis Pitas, and Nikos Nikolaidis. 2001. Robust audio watermarking in the time domain. *IEEE Transactions on multimedia* 3, 2 (2001), 232–241.

[4] Walter Bender, Daniel Gruhl, Norishige Morimoto, and Anthony Lu. 1996. Techniques for data hiding. *IBM systems journal* 35, 3.4 (1996), 313–336.

[5] Dongjian Cai and Kaliappan Gopalan. 2014. Audio watermarking using bit modification of voiced or unvoiced segments. In *IEEE International Conference on Electro/Information Technology*. IEEE, 491–494.

[6] M Chetan, Prarthana P Bhat, Vrushabh Shet, Sana Begum Husenbhai, and Ashwini Bhat. 2021. Audio Watermarking Using Modified Least Significant Bit Technique. In *2021 International Conference on Circuits, Controls and Communications (CCUBE)*. IEEE, 1–5.

[7] Maria Chroni, Angelos Fylakis, and Stavros D Nikolopoulos. 2014. From Image to Audio Watermarking Using Self-Inverting Permutations.. In *WEBIST (1)*. 177–184.

[8] Maria Chroni, Stavros D Nikolopoulos, and Leonidas Palios. 2018. Encoding watermark numbers as reducible permutation graphs using self-inverting permutations. *Discrete Applied Mathematics* 250 (2018), 145–164.

[9] Nedeljko Cvejic and Tapio Seppanen. 2002. Increasing the capacity of LSB-based audio steganography. In *2002 IEEE Workshop on Multimedia Signal Processing*. IEEE, 336–338.

[10] Yousof Erfani and Shadi Siahpoush. 2009. Robust audio watermarking using improved TS echo hiding. *Digital Signal Processing* 19, 5 (2009), 809–814.

[11] Kaliappan Gopalan and Qidong Shi. 2010. Audio steganography using bit modification-A tradeoff on perceptibility and data robustness for large payload audio embedding. In *2010 Proceedings of 19th International Conference on Computer Communications and Networks*. IEEE, 1–6.

[12] Guang Hua, Jonathan Goh, and Vrizlynn LL Thing. 2015. Cepstral analysis for the application of echo-based audio watermark detection. *IEEE Transactions on Information Forensics and Security* 10, 9 (2015), 1850–1861.

[13] Guang Hua, Jiwu Huang, Yun Q Shi, Jonathan Goh, and Vrizlynn LL Thing. 2016. Twenty years of digital audio watermarking—a comprehensive review. *Signal processing* 128 (2016), 222–242.

[14] Md Islam, Nuzhat Naqvi, Aliya Tabassum Abbasi, Md Hossain, Rizwan Ullah, Rashid Khan, M Shujah Islam, Zhongfu Ye, et al. 2021. Robust dual domain twofold encrypted image-in-audio watermarking based on SVD. *Circuits, Systems, and Signal Processing* 40, 9 (2021), 4651–4685.

[15] Aniruddha Kanhe, Gnanasekaran Aghila, Ch Yaswanth Sai Kiran, Ch Hanuma Ramesh, Gabbar Jadav, and M Gowtham Raj. 2015. Robust audio steganography based on advanced encryption standards in temporal domain. In *2015 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*. IEEE, 1449–1453.

[16] Wen-Nung Lie and Li-Chun Chang. 2006. Robust and high-quality time-domain audio watermarking based on low-frequency amplitude modification. *IEEE transactions on multimedia* 8, 1 (2006), 46–59.

[17] Noureddine Mehallegue, Khaled Loukhaoukha, Khalil Zebbiche, Ahmed Refaey, and Mourad Djellab. 2020. Ambiguity attacks on SVD audio watermarking approach using chaotic encrypted images. *Multimedia Tools and Applications* 79, 3 (2020), 2031–2045.

[18] Mohamed Yamni, Hicham Karmouni, Mhamed Sayyouri, and Hassan Qjidaa. 2022. Efficient watermarking algorithm for digital audio/speech signal. *Digital Signal Processing* 120 (2022), 103251.