The International Journal of Engineering and Information Technology

journal homepage:www.ijeit.misuratau.edu.ly

Copyright Protection Based on Hybrid Image Watermark (DCT) with Audio Watermark (EMD)

AL Darrat, k.
Computer Science Department,
Misurata University
k.ali@it.misuratau.edu.ly

Abushaala, A.
Computer Science Department,
Misurata University
a.mohamed@it.misuratau.edu.ly

Abstract - **A digital watermark is a type of security that incorporates data into a signal, making it difficult to erase and showing ownership and validity. An important issue brought about by the growing use of the Internet is illicit copyright infringement and authentication. Previous research has proposed several techniques for this problem but has proven them to be either robust, difficult to detect, or fragile and easy to break. Because of this, in this paper, we use two separate kinds of watermarks—one for image and one for audio—to merge these two purposes into one. The image watermark discrete cosine transform (DCT) is renowned for its robustness. It is embedded in an audio watermarking algorithm based on empirical mode decomposition (EMD), ensuring a safe and fragile embedding process. The audio signal was divided into frames, and each one was decomposed by EMD into intrinsic mode functions (IMFs). The DCT watermark bits and the synchronization codes are embedded into the extrema of the first IMF, a high frequency mode sensitive to any small changes in the audio signal. The experimental outcomes suggest that while using two separate kinds of watermarks may seem like a comprehensive solution, it shows the flexibility of EMD and the limits of the DCT watermark technique in different scenarios.**

Keywords: **audio signal, EMD, watermark, robust, fragile, copyright, DCT, authentication, algorithm.**

I. INTRODUCTION

Digital audio watermarking has attracted a lot of attention as effectively protecting digital media. This technique involves storing or embedding a watermark into the original audio stream—which can be images, text, or sound. for use in a variety of applications including broadcast monitoring and copyright protection (Cano et al. 2005). A popular study area right now is content validation(Anderson 2020). Producers can protect their work from unauthorized use or distribution by incorporating watermarks into audio streams.

These technologies are critical to maintaining the integrity and authenticity of digital content in today's fast-paced media environment. The watermark enables you to authenticate the audio source by detecting any unauthorized changes or modifications and ensures data integrity. These policies protect content creators' intellectual property rights and help identify and prevent unauthorized copying and distribution activities The rise of digital transformation has made the inclusion of watermarks an important tool to protect the integrity and authenticity of digital information Who will (Boney, Tewfik, and Hamdy 1996).

Digital audio watermarks are classified into three types based on their specific purpose: fragile, semifragile, and robust watermarks (Wang and Fan 2010) Fragile watermarks are very sensitive and recognizable there are slight changes to the original multimedia file, making them useful for accurate approximate integrity. A semi-fragile watermark may cope with some normal signaling behavior but may not prevent malignant changes. Robust watermarks, as the name suggests, are primarily used for copyright protection because they are highly resistant to various signaling attacks.

In the case of digital watermarks, it is important that the watermark used is robust, imperceptible, and has data capabilities efficient enough to identify the owner of the media If data integrity is to be maintained, the watermark must be invisible in the image or audio data. It is also important that the watermark is easily extracted to provide proof of ownership. Therefore, digital watermarks should be invisible and easy to (Khaldi, Boudraa, and processing 2012).

The hybrid watermark approach is a powerful tool for enhancing the security of digital media. These techniques combine digital and analog features, making it difficult to remove a watermark from the device during a signal attack. In addition, the hybridization process ensures that watermarks remain invisible and invisible to the human eye. This is important to prevent unauthorized use of copyrighted material (Patel, Parikh, and Applications 2023).These strategies are essential to protect the rights of creators in the digital era.

Received 30 Apr , 2024; revised 10 May, 2024; accepted 15 Mar 2024.
AVAILABLE ONLINE 08 AUG, 2024.

II. OVERVIEW OF THE PROPOSED METHOD

A. *Synchronization code*

We used SCs to locate the embedding position. Let A be the original SC and B be an unknown sequence of the same length. We consider sequence B as an SC if only the difference between A and B is less than or equal to a predefined integer threshold τ between (0, 10) (Abdulmunem and Badr 2017).

B. *Quantization index modulation (QIM)*

QIM has lately become a common approach to watermarking based on the framework of communication with side information. We pick our approach (QIM) owing to its great robustness and blind nature. The settings of QIM are changed to guarantee that the embedded watermark in the last/first IMFs is inaudible. The watermark is linked with a synchronization code to facilitate its localization. A benefit of adopting the time domain strategy, based on EMD, is the affordable cost of hunting for synchronization codes. This is because the synchronization codes are immediately included in the transmission without any additional processing.

C. *Quick Response (QR) codes*

The QR code is a two-dimensional symbol. It was invented in 1994 by Denso. It can store large amounts of alphanumeric information and is easily readable by a scanner. QR codes are interference-free because of their 360-degree readability, and their data storage capacity, which is determined by their data-type mode. One can see a QR code with different versions in FIG.1.

FIG. 1 QR code versions: (a) Version 1 (21×21). (b) Version 4 (33×33). (c) Version 10 (57×57). (d) Version 40.(177×177)

D. *Proposed Approach*

The Proposed approach goes through several steps, as shown in FIG.2

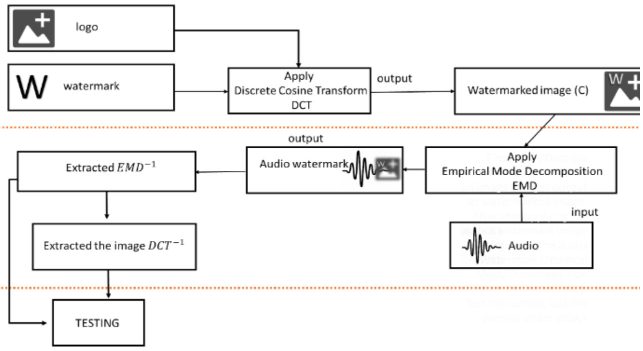


FIG. 2 Proposed Approach

The discrete cosine transform (DCT) is a widely used technique in the transform domain. In DCT-based watermarking methods, the host image is first divided into non-overlapping blocks of size 8x8. Then, the DCT is applied to each block to obtain the corresponding DCT coefficients. These coefficients are categorized as low-frequency (LF), mid-frequency (MF), and high-frequency (HF) based on their frequencies. The first low-frequency coefficient is known as the direct-current (DC) coefficient. FIG. 3 illustrates these coefficients using different color codes for better understanding. (Sharma et al. 2024)

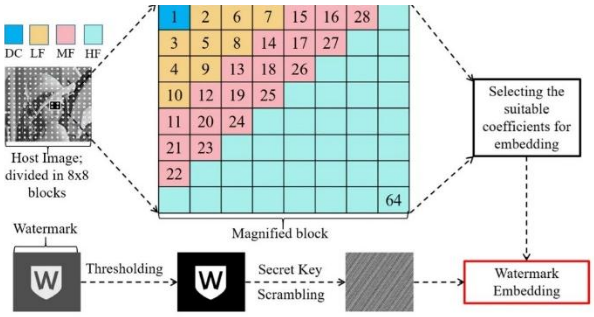


FIG. 3Discrete Cosine Transformation

Empirical Mode Decomposition (EMD) is a non-linear, data-driven signal processing method that decomposes a signal into its intrinsic mode functions (IMFs). In EMD, the audio signal is decomposed into different intrinsic mode functions (IMFs) by segmenting it. The watermark is then applied to the IMFs as binary bits using Quantization Index Modulation (QIM).

This paper proposes a watermark that combines fragility and robustness simultaneously, based on robust and secure image watermarking via DCT for copyright protection embedded in adaptive, fragile audio watermarking for content authentication via EMD.

III. WATERMARKING PHASES

- **Phase One:** It is a well-known mathematical tool used to transform an image from the spatial domain to its frequency domain. It relates to the mathematical equation (Eq. 1).
  A non-overlapping selected 8 × 8 block of the cover image is used to hide the 1-bit information of the watermark image (binary image). A selected mid-frequency coefficient of the DCT position is used here to hide the information. The total watermark information (bit pattern) is embedded in the same manner by selecting the different 8 × 8 blocks of the cover image. After hiding each bit, the watermarked image is reconstructed by applying IDCT to each 8 × 8 block (Liu, Pan, and Song 2017).

$$\mathrm{DCT}(u) = c(u) \cdot \sum_{i=1}^{S} \mathrm{IDCT}(i) \cdot \cos\left(\frac{i+(1/2)}{S} \cdot u\pi\right) \quad (1)$$

Step 1: When we receive the host image and logo, convert the image to a grayscale and resize it to 512*512. Moreover, reformulate the logo into a sequence of 1 and 0 (black and white) and resize it to 32*32.

Step 2: Splits the host image into non-overlapping blocks of 8×8 pixels.

Step3: For each block transformed by DCT, the image blocks with the highest variance value were selected for embedding the watermark. The purpose of selecting an image block with high variance pixels is to get an image block that is less sensitive to the human eyes. The selected image blocks considered the watermark size. Where number of selected image blocks is equal to 1024 image blocks because the experiments used a binary watermark with the size of 32 × 32 pixels .

Step 4: Recompose the image by inverting DCT coefficients of the image blocks that contain the watermark.

- **Phase two**: the input audio signal is segmented into equal frames, and EMD is conducted on every frame to extract the associated IMFs (see FIG.4). The number of IMFs depends on the amount of data in the frame, so the number of watermark bits to be added is different in each frame. It depends on the length of the frame too.

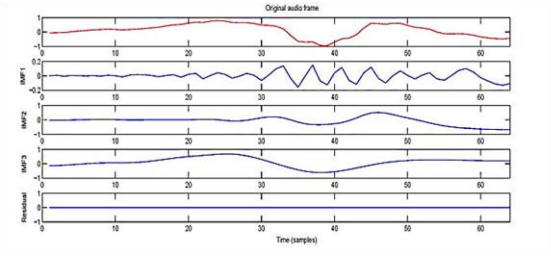


FIG. 4 Decomposition of an audio frame by EMD

Before embedding, convert the watermark images to black and white, and then create a binary data sequence consisting of SCs and informative watermark bits (see FIG.5) embedded in the extrema of a set of consecutive first-IMFs. (zero or 1) inserted per extrema.

Sync-code	Watermark bit	Sync-code

FIG. 5 sequence consisted of SCs and informative watermark bits

Watermark data and SCs are not all embedded in the extrema of the first IMF of only one frame. Compared to the length of the binary sequence to be embedded, the number of extrema per first IMF in one frame is very small.

Assuming we design by L1 and L2 the numbers of bits of SC and watermark, respectively, the length of the binary sequence to be embedded is equal to 2L1 + L2. Thus, these 2L1+L2 bits spread out on several first-IMFs (extrema) of the consecutive frames. Further, this sequence of 2L1+L2 bits embeds P times. Watermarking embedding processes are summarized in FIG.6.

Finally, the modified extrema of the watermarked signal was transformed back by using (EMD^{-1}), followed by frame concatenation.

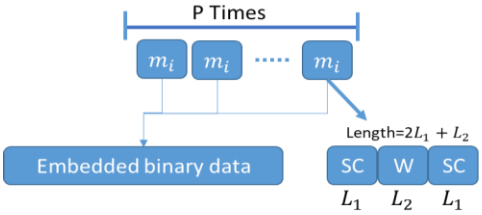


FIG. 6 Data extraction

For data extraction, the watermarked audio signal is split into frames, and EMD is applied to each frame (see FIG.5) to get IMFs. Searching for SCs from each first IMF to extract binary data sequences.

We should know the watermarking technique is hidden; if the number of IMFs is not the same, there is no assurance that the first IMF will contain the necessary data for watermark extraction. Consequently, the watermark must be able to extract the same quantity of IMFs and length of frames as it did before watermarking.

IV. WATERMARK EMBEDDING

Before embedding, it combined SCs with watermarked images to form a binary sequence denoted by M(i) ∈ [1], where (i) is the watermark bits as seen in FIG .5.

Step 1: Consists of segmenting the audio signal into frames.

Step 2: Each frame is decomposed by EMD into IMFs as follows (Eq. 2):

$$x(t) = \sum_{j=1}^{C} \mathrm{IMF}_j(t) + \mathrm{r}_C(t) \tag{2}$$

Where **x (t)** is the signal, **c** is the number of IMFs, and **rC (t)** denotes the final residual.

EMD is a fully data-driven method that recursively breaks down any signal into a reduced number of zero-means with symmetric envelope AM-FM components called Intrinsic Mode Functions (IMFs). Where bits are inserted into the extrema of the first IMFs, such that the watermarked signal inaudibility is guaranteed.

Step 3: Embed p times the binary sequence [1] into the extrema of the first IMFs by QIM (Eq. 3):

$$e_i^* = \begin{cases} \lfloor e_i/\mathrm{S} \rfloor \cdot \mathrm{S} + \mathrm{sgn}(3\mathrm{S}/4) & \text{if } m_i = 1 \\ \lfloor e_i/\mathrm{S} \rfloor \cdot \mathrm{S} + \mathrm{sgn}(\mathrm{S}/4) & \text{if } m_i = 0 \end{cases} \tag{3}$$

Where ei and ei* are the extrema of IMFs, the host audio signal, and the watermarked signal, respectively. The sgn function is equal to "+" if ei is maxima, and "-" if it is minima. []denotes the floor function, and S denotes the embedding strength chosen to maintain the inaudibility constraint.

Step 4: Reconstruct the frame (EMD-1) using modified IMFS and concatenate the watermarked frames to retrieve the watermarked signal.

V. PERFORMANCE ANALYSIS

Imperceptibility is our primary objective; the audio signal quality should not degrade after adding the watermark. Imperceptibility can be assessed using an objective measure, such SNR, which should be more than 20 db. According to IFPI recommendations.

The performance of our method in terms of the objective test for DCT was measured by PSNR between the original and the watermarked image, as seen in equation (Eq 4) before embedded in EMD and after. The data payload of our method is also a crucial factor in determining its performance.

Furthermore, the signal-to-noise ratio (SNR) between the original and watermarked audio signals is an important metric for evaluating the effectiveness of the watermarking process, as seen in equation (Eq 5). According to the International Federation of the Photographic Industry (IFPI), a watermark audio signal should maintain more than 20 dB SNR.

$$PSNR = 10\log_{10}\left(\frac{MAX_i^2}{MSE}\right) \qquad (4)$$

MAX is the maximum pixel value of the image. When the pixels are represented using 8 bits per sample, this is 255. Mean squared error (MSE).

$$\mathrm{SNR}(\mathrm{A},\mathrm{A}^{\sim}) = 10\,\mathrm{log10}\,\frac{\sum_{n=1}^{N} A^2\,(\mathrm{n})}{\sum_{\mathrm{n}=1}^{\mathrm{N}}(\mathrm{A(n)}-\mathrm{A}^{\sim}(\mathrm{n}))^2}\ \mathrm{db} \qquad (5)$$

The audio signal A represents an N-sample audio signal. While A˜ is the watermarked version of that signal,

In addition to identifying the two QR codes after recovering them from the audio and image watermarks.

QR1

QR2

Audio File	Original	SNR (dB)	Image-extract	Can be decoded?
Classic music		39.2		Yes
Latin music		42.6		Yes
Pop song		37.358		Yes
Slow song		33.6		Yes

Table.1

We used two QRs one as an image QR1 (Content is Authenticate) and the other QR2 as a logo (DCT).

VI. RESULT

To show the effectiveness of our proposal, we perform simulations on audio signals. The embedded watermark, W, is a binary logo that we saved before by DCT and resizes the image to M*N =34×34 bits (FIG. 7). Convert this 2D binary image into a 1D sequence to embed it into the audio signal in wave format sampled at 44.1 kHz with 16-bit depth, and a 10-second duration for each audio signal. We divide all audio signals into frames of size 64.

FIG. 7 Binary watermark

FIG.8 shows a portion of the signal and its watermarked version. This FIGURE shows that the watermarked signal is visually indistinguishable from the original one.

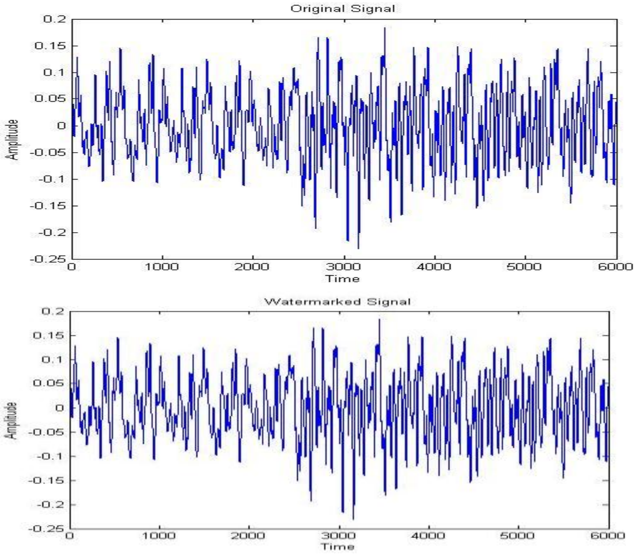


FIG. 8 portion of the signal and its watermarked version

Table.1 presents the results of perceptual quality tests where the signal-to-noise ratio (SNR) was measured. It is important to note that the SNR values are not equal because the EMD method used is data-dependent. However, all QR codes extracted from the watermarked audio are readable. Our results indicate that our method is effective in maintaining an SNR of over 25 dB for all audio signals tested.

Reviewing Table1. EMD managed the watermarked image the same way it managed a regular image, with only a minor change to the PSNR values, and the watermark was successfully retrieved. While the DCT failed to extract the logo from the image, a PSNR evaluation was not necessary to verify this Table.2 Showing the retrieve the logo with DCT from the extracted watermark and work out the PSNR.

Audio File	Original	PSNR	Extracted logo After EMD	Can be recognized
Classic music		3.21		NO
Latin music		3.20		NO
Pop song		3.2		NO
Slow song		3.43		NO

Table.2

The results from the experiment were different from what was expected. Therefore, we had to modify a few things, check for an enhancement in the results, and discover the explanation for this failure.

- First, we started with some experiments on the image algorithm. We switched the image type from gray to black and white, and then we checked to see if DCT could identify the logo (QR) in Table.3

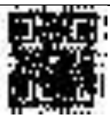
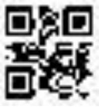
Image type BMP	watermarked jpg image	Logo	Extracted logo	PSNR
16 Bit				18.9195
24 Bit				41.5637

Table.3 DCT PSNR for different Bits

- Second, we tried adjusting the image size and experimented with different image sizes to see if that improved the result Table.4

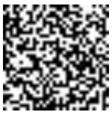
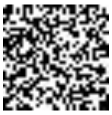
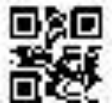
Watermark Image size	Extracted logo	PSNR	Can be recognized
34x34		3.9319	NO
100x100		13.7592	NO
250x250		21.8007	NO
500x500		29.8728	NO
512x512		41.5638	Yas

Table.4 DCT PSNR for different sizes

- At last, we found that the image that the audio algorithm handles is tiny, with a size of 34 * 34. We increased the size of the input image by increasing the audio sample size, raising the number of IMFs to 1051451 bits by extending the time from 10 seconds to 6 minutes.As we show early in each IMF we can add just one bit of image. The results were Table.5.

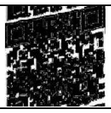
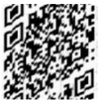
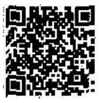
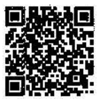
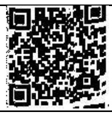
Audio Size	Size	Extrema.nu	SNR (dB)	Image-extract	PSNR	Can be recognized
1 min	512*512	96004	Not fit	----	---	NO
3 min	512*512	753586	50.66		3.7241	NO
6 min	512*512	1051451	80.24		4.3050	NO
1 min	100*100	96004	50.65		21.810	YES
3 min	100*100	753586	14.92		22.1796	YES
6 min	100*100	1051451	12.64		19.6783	NO

Table.5

VII. DISCUSSION

We have tried everything to get the audio watermark to function, but to no avail. The failure was not a consequence of EMD since it was able to get the QR logo in two distinct formats and sizes. It was the DCT watermark itself that caused the issue. The conversion from grayscale to black and white was not the major issue, but the resizing was denied via the DCT. Since DCT decomposes images into spatial frequency domains, the image is ordered into a predetermined spatial course. If the image dimensions affect the spatial arrangement of pixels, the DCT frequency components may not be a fair depiction of the original image. Because of this, the rebuilt image may suffer from compression imbalance, capabilities are lost, and pleasure is lowered. Sizing increases discrepancies in mass boundaries, limiting the possibility of proper recovery.

VIII. CONCLUSION

Our observation led us to conclude that EMD was the right choice for our proposal, as it sensitively (fragile) and efficiently preserved audio without being affected by the other watermark, successfully achieving the first feature. Assuming that the computer had sufficient RAM and that the picture size matched the number of IMFs extracted from the audio, this audio technique could handle watermarks of any size, even enormous ones. It also yields a high SNR of more than 50. The audio quality persisted, but at 512 x 512 pixels, the recovered image was not too terrible despite being widely dispersed over the 262144 IMFs utilized in the audio.

We also noticed that Even though DCT has a great ability to retrieve images and PSNR>40 was excellent in normal conditions, it failed when combined with the audio watermark. Yet, like any other technique, the DCT image watermarking technique has limitations. One of the main limitations is the sensitivity to geometric attacks. It is not capable of handling the size and type changes necessary to be embedded in the other watermark, therefore using it for our proposal won't work since it can't handle the size and type modifications required to be embedded in the EMD watermark.

DCT was unable to recover the image due to bit loss and differences in the row and column addresses for the matrix on each block, which makes it impossible to restore the logo.

However, like any other technique, the DCT image watermarking technique also has its limitations. One of the main limitations is the sensitivity to geometric attacks. If the watermarked image is rotated, scaled, or translated, the watermark may not be accurately extracted, resulting in errors.

IX. FUTURE WORK

For future work, When considering combining two watermarks, it's important to take into account not just the robustness of the second watermark but also its limitations and vulnerabilities before starting the merging process to prevent the first watermark from seeing the second as an attack and being destroyed. In this instance, selecting a watermark that is resilient to engineering attacks, one that can accept change—is necessary(Cut and resize), opposite from DCT.

REFERENCES

1. Abdulmunem, Matheel E, and Ameer A %J International Journal of Advanced Research in Computer Science Badr. 2017. 'Fragile Audio Watermark based on Empirical Mode Decomposition for Content Authentication', 8.
2. Anderson, Ross. 2020. *Security engineering: a guide to building dependable distributed systems* (John Wiley & Sons).
3. Boney, Laurence, Ahmed H Tewfik, and Khaled N Hamdy. 1996. "Digital watermarks for audio signals." In *Proceedings of the third IEEE international conference on multimedia computing and systems,* 473-80. IEEE.
4. Cano, Pedro, Eloi Batlle, Emilia Gómez, Leandro de CT Gomes, Madeleine %J Computational intelligence for modelling Bonnet, and prediction. 2005. 'Audio fingerprinting: concepts and applications': 233-45.
5. Khaldi, Kais, Abdel-Ouahab %J IEEE transactions on audio Boudraa, speech,, and language processing. 2012. 'Audio watermarking via EMD', 21: 675-80.
6. Liu, Shuai, Zheng Pan, and Houbing %J IET image processing Song. 2017. 'Digital image watermarking method based on DCT and fractal encoding', 11: 815-21.
7. Patel, Swati J, Mehul C %J Reliability: Theory Parikh, and Applications. 2023. 'HYBRID AND BLIND WATERMARKING FRAMEWORK FOR PRIVACY PROTECTION AND CONTENT AUTHENTICATION OF DIGITAL MULTIMEDIA', 18: 456-65.
8. Sharma, Sunpreet, Ju Jia Zou, Gu Fang, Pancham Shukla, Weidong %J Multimedia Tools Cai, and Applications. 2024. 'A review of image watermarking for identity protection and verification', 83: 31829-91.
9. Wang, HongXia, and MingQuan %J Science China Information Sciences Fan. 2010. 'Centroid-based semi-fragile audio watermarking in hybrid domain', 53: 619-33.