E. Duymaz and A. Akan/ IU-JEEE Vol. 12(2), (2012), 1517-1522

NOVEL DIGITAL AUDIO WATERMARKING APPROACH

Erol DUYMAZ[1], Aydın AKAN [2]

[1]Elektronik Mühendisliği, Hava Harp Okulu, İstanbul
[2]Elektrik- Elektronik Mühendisliği, İstanbul Üniversitesi – İstanbul
e.duymaz@hho.edu.tr, akan@istanbul.edu.tr

Abstract: *Watermarking is a basic secure communication method. It is used for embedding a recognizable pattern in media in such a manner that modification of the media also modifies the pattern, thus making it easy to detect the modification. This technique and its variants have many practical applications pertaining to secure communications, media verification, etc. Digital audio watermarking is a technique for embedding data within an audio signal in such a way that the original and the modified audio signals are essentially identical. The embedded data can be used for various purposes such as secure communication in military applications, owner identification and verification, content authentication, etc. In this study a watermark audio signal is hidden in a message audio signal by using a discrete cosine transform (DCT) domain approach. The tests of fidelity between the original and the watermarked signal and robustness applied to the watermarked signal. The results with both the Human Auditory System (HAS) and numeric/graphic aspects are presented. The results show that an embedded watermark is not easily detectable using either the HAS or other techniques. Additionally, it can be detected successfully in the simulation domain, but it may be susceptible to some noise and channel limitations in the real world.etc.*
Keywords: *Discrete Cosine Transform, audio watermarking.*

1. Introduction

Communication between members of a species has existed since the advent of the species. Leaping ahead to the human species, communication started when some human needed to convey some information to another and understand them in return. Since those early forms of communications in gestures and sounds, much has changed in both the style and methods of communication. Nowadays humans communicate with each other using multiple techniques and methods not limited to face-to-face verbal communication. Today it is possible, for instance, to communicate with people a long distance away over a wire or through a small device held in your hand.

Sometimes people need to hide their communication from everyone except the one intended to receive the communication. For instance, a prisoner who wants to plan an escape may try to hide his communications with his conspirator using some kind of code, or a battle commander needs to hide his message to officers in the field using some other technique. So the need for ways to perform secure or unshared communication arose.

On the other hands with the spread of the Internet in recent years, digital multimedia works like video, audio and images have become increasingly available for electronic transmission, production, and publishing. Connected to this increase in the use of digital media is the strong desire for protection against unauthorized copy and propagation to protect financial and proprietary rights. These concerns triggered research for finding ways to deter copyright trespassing. One of the best solutions for this challenging problem looks to lie in information hiding techniques. Information hiding is the process of embedding a message into the digital signal. The embedded message needs to be audibly imperceptible.

Watermarking embeds a recognizable pattern in transmission media to provide authentication. Digital audio watermark is a technique for embedding additional data along with the audio signal. Embedded data is used for various purposes such as copyright protection, owner identification, or security.

A number of audio watermarking techniques that has increased sharply in recent years to embed a robust watermark and keep original signal fidelity are in existence today[1]. Alsalamai and Al-Akaidi [2] present a good survey of digital audio watermarking principles, and Kim [3], Lee and Jung [4], Xu at all [5] examined digital audio watermarking techniques in detail and Chenga at all [6] compared performances of watermarking algorithms in their studies. Among different methods Wang and Zhao [7] embeds a watermarking signal in low frequency coefficients of discrete wavelet transform (DWT), Bath at all [8] in cepstrum transform coefficients by using quantization, Ramalingam and Krishnan [9] by using short-time fourier transform (STFT) for audio fingerprinting and Megias at all [10] by using fast fourier transform (FFT).

Although there are many digital watermarking algorithms in the literature today, generally speaking, each has some drawbacks even while it is sufficient in

Received on: 19.03.2012
Accepted on: 13.06.2012

other aspects. This is the reason why researchers continue to look for better algorithms. In this proposed novel algorithm that is applicable to many applications ranging from multimedia to military uses and simple to implement, a watermark audio signal will be hidden in a message signal in the discrete cosine transform (DCT) domain and tests will be performed on the watermarked signal to verify robustness and fidelity then the algorithm is evaluated thoroughly by different host signals from a human voice to music or by the effect of noise, bandwidth limitations and crop attacks in a simulation environment and discussed by its pros and cons versus other approaches existent in the literature.

2. Method Application and Results

Generally a watermarking system consists of three modules; watermark signal generation module, watermark embedding module and watermark detection module [3]. The watermark signal is generated using a non-invertible function that takes as an input a watermark key. In some systems the host signal (cover-object) is taken into account when the watermark is generated. This will help the watermark generator in producing an imperceptible signal-dependent watermark [3]. Watermark embedding can be performed either in the time domain or in the transform domain (DFT, DCT, DWT, etc.) using a suitable embedding rule (e.g., addition or multiplication). Finally, watermark detection is performed by some sort of correlation detector or statistical hypothesis testing, with or without resorting to the original signal [3].

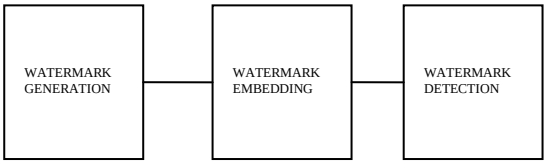

Figure 1. A general watermarking system

The main issue in watermarking schemes is to decide how to embed the watermark into the message. Today many ways exist in the literature. While the nonblind algorithms tend to be very robust, their requirement that the original signal is needed to detect the watermark is not always practical. For this reason, blind approaches, which do not require the original signal to detect the watermark signal, are preferred among researchers.

The watermark signal that is recorded in a computer and shown in time domain in Figure 2 is to be embedded into the the original audio message which is to be transmitted. By using the DCT, the two signals are transformed to the frequency. Here the signals can be represented by the coefficients of DCT as shown in Figure 3.

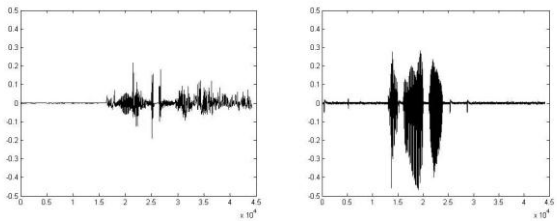

Figure 2. Message signal (left) and watermark signal (right) in time domain.

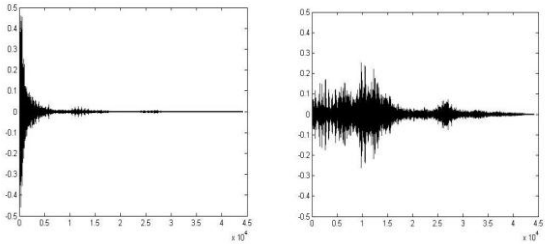

Figure 3. Magnitude of the DCT coefficients of the message (left) and those of the watermark signal (right).

The main idea is to embed the DCT coefficients of the watermark signal into the coefficient of the message signal. The process for embedding the watermark into the signal is fairly straightforward and relies on quantizing the DCT coefficients of the two signals. The floating-point representation of a DCT coefficient is represented by:

$$c = d_n d_{n-1} \ldots d_0 \cdot m_1 m_2 \ldots m_p \qquad (1)$$

where c is the DCT coefficient, d_i, i = 1,...n are the digits to the left of the decimal and m_k, k = 1,...,p are the mantissa, or the digits to the right of the decimal sign. The algorithm for data embedding is as follows:

1. Truncate the DCT coefficient of the message to 4 decimal places, i.e., p = 4.
2. Truncate the DCT coefficient of the watermark to 4 significant digits, i.e., p = 4.
3. Multiply the truncated watermark coefficient by 10^{-5} to shift it 5 places to the left. After the multiplication, the watermark DCT coefficient will take the form: 0.0000.....
4. Concatenate the truncated signal coefficient with the shifted truncated watermark coefficient to form the DCT coefficient of the embedded message.
5. Take the inverse DCT of the concatenated signals, and transmit.

This procedure forms a DCT combined DCT coefficient that has 10 digits in the mantissa. An 11th digit needs to be added to the representation to indicate the sign of the coefficient of the watermark signal. If the watermark signal coefficient and the message signal coefficients have the same sign, then the sign digit is set to zero. If they differ then the sign digit is set to 1 and the sign of the watermark coefficient—the signal to be retrieved—is obtained from the sign of the

DCT coefficient of the transmitted data. If the coefficient of the transmitted signal is negative and the sign digit is a 1, the DCT coefficient of the watermark is positive and vice versa.

Let us use an example to clarify the process. The first DCT coefficient of the message signal is 3.1234567, and the first DCT coefficient of the watermark signal is 0.9876543. Since the coefficient does not change significantly after the 4th digit in the decimal fraction let us truncate digits after 4. Then, the first DCT coefficients of the message and watermark signals are, respectively, 3.1234 and 0.9876. Now the DCT coefficient of the watermark signal is shifted to the right by 5 digits which makes it 0.000009876. Concatenation is then a simple matter of adding the two coefficients, giving the new coefficient as: 3.1234000000 + 0.000009876 = 3.123409876. Here it is obvious that 3.1234 is derived from the first DCT coefficient of the message signal, and the rest, 09876, is derived from the first DCT coefficient of the watermark signal.

The last step is to assign the last digit associated with the sign of the DCT coefficient of the watermark. There are 4 possible cases for the DCT coefficients of the message and watermark signals: + + , - - , + - , and - + , where the '+' and '-' indicate the sign of the coefficients of the two signals. Since the watermark signal coefficients are concatenated with the message signal coefficients by shifting and addition, the sign of the watermark coefficient is needed to perform the correct addition operation. If the signs of the two coefficients are different, the addition operation would change the magnitude of the coefficient of the DCT coefficient of the original signal rather than just concatenating it. Hence, the sign bit needs to be appended to the transmission coefficent to ensure that the correct information is decoded at the receiver. For example, when the DCT coefficient of the message signal is 3.1234567 and the DCT coefficient of the watermark signal is -0.9876543, the truncation, shift and add processes would produce: 3.1234000000 + (-0.000009876) = 3.123390124 which would be interpreted as the DCT coefficient of the original signal, 3.1233, and the DCT coefficient of the watermark signal is interpreted as 9.0124 which is incorrect. Hence, the addition is performed as: 3.1234000000 + abs(-0.000009876), which produces the correct result, and a sign digit of 1 is appended in the leftmost position to indicate the change in sign between the two coefficients.

The experimental results are given in Table 1. These are the numeric values of the first 20 DCT coefficients of the original (left) and the watermark data (right). However, the key comparison between the message signal and the watermarked signal does not show perceptible visual or audible changes as shown in Table 2. At the decoder, the last digit determines the sign of the watermark coefficient. Look at the example and again let the DCT coefficient of the message signal be 3.1234567 and the DCT coefficient of the watermark signal be -0.9876543.

Table 1. The message and the watermark signal DCT coefficients respectively.

DCT Coefficients	
Message Signal	Watermark Signal
0,092584	0,101460
0,011473	6,8007510e-05
-0,011769	-0,000979
0,008292	-0,000954
-1,259911e-05	0,005661
-0,006420	0,000532
0,001601	0,002223
0,018989	-0,000898
-0,025747	-0,001886
0,047255	-0,004458
-0,049662	0,003798
0,024227	-0,006615
-0,026798	0,006007
0,057598	0,009358
-0,028952	0,003933
-0,036668	0,002528
0,036598	-0,003824
-0,048740	-0,010114
0,110783	0,009156
0,004680	0,004372

Table 2. The DCT coefficients of the retrieved message (left) and watermarked signal (right).

DCT Coefficients	
Retrieved Message Signal	Watermarked Signal
0,092584	0,09201010000
0,011473	0,01100000000
-0,011769	-0,01100000000
0,008292	0,008000010000
-1,259911e-05	-5,100000000e-07
-0,006420	-0,006000010000
0,001601	0,001000200000
0,018989	0,01800001000
-0,025747	-0,02500010000
0,047255	0,04700041000
-0,049662	-0,04900031000
0,024227	0,02400061000
-0,026798	-0,02600061000
0,057598	0,05700090000
-0,028952	-0,02800031000
-0,036668	-0,03600021000
0,036598	0,03600031000
-0,048740	-0,04800100000
0,110783	0,1100009000
0,004680	0,00400040000

The respective DCT coefficient of the combined signal at the receiver is 3.1234098761. Here the last digit of 1 indicates that the sign of the watermark coefficient is opposite to the sign of the message coefficient, so it is negative. This means that during

the detection operation, the value 09876 is to be read as -0.9876.

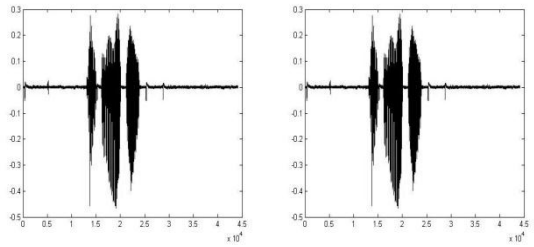

Figure 4. The detected watermark signal (right) and the orginal watermark signal (left) in time domain.

This procedure was used in the algorithm. After conducting tests it is observed that the watermark signal was detected and retrieved successfully. Figure 4 depicts the retrieved watermark to compare this with the original watermark signal. It is very difficult to visually compare the two figures and see any significant differences.

The DCT coefficients of the original watermark signal and detected watermark signal and their difference are given in Table 3. The magnitude of DCT coefficients of the original watermark and the detected watermark signal are also shown in Figure 5. Errors due to truncation are evident.

Table 3. The DCT coefficients of the original watermark signal (left) , the detected watermark signal (middle) and their differences (right).

DCT Coefficients		
Watermark Signal	Retrieved Watermark Signal	Difference
0.101460678	0.101	0.000461
0.000006800	0.000	0.000068
-0.000979158	0.000	-0.000980
-0.000954252	0.000	-0.00095
0.005661341	0.005	0.010661
0.000532853	0.000	0.000533
0.00222315	0.002	0.000223
-0.000898385	0.000	-0.00090
-0.001886812	-0.001	-0.00089
-0.004458484	-0.004	-0.00846
0.003798939	0.003	0.006799
-0.006615473	-0.006	-0.01262
0.00600784	0.006	0.012008
0.009358237	0.009	0.000358
0.003933716	0.003	0.006934

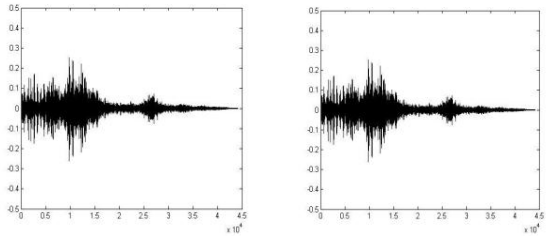

Figure 5. Magnitude of the detected watermark signal DCT coefficients (left) and those of the original watermark signal.

4. Transmission Considerations

The results in previous part show numerically and graphically that a watermark signal can be embedded in a message signal using the proposed method and be detected and retrieved succesfully. These results were also tested by listening to the original watermark and the retrieved watermark. However, the original algorithm testing was performed in a clean simulation environment. There was no signal fading or channel noise that could impact the watermarked signal and, hence, its DCT coefficients. In this part some issues related to real-world channels are examined. The communication toolbox in Matlab is used for these tests.

The main idea is that in a communication system, when the transmitter side sends its message to the receiver side through a wired medium, the signal experiences some losses due to the length and the tranmission quality of the cable, or, if it is wireless transmission, some losses due to weather effects. Since the high frequency components dominantly characterize an audio signal, these components should be examined for both cropping and bandwidth effects.

The channel limitations are one of the major considerations and are simulated by a low pass filter. When the watermarked signal is filtered by a Butterworth lowpass filter,

$$H(z)=\frac{b(1)+b(2)z^{-1}+....b(n+1)z^{-n}}{a(1)+a(2)z^{-1}+...a(n+1)z^{-n}} \qquad (2)$$

where b and a are length n+1 row vectors that represent the filter coefficients in descending powers of z. It is observed that the watermark is detectable only if the cut-off frequency ω_C of the filter is $0.01F_s$, where F_s is the sampling frequency of the signal. This cut-off frequency corresponds to about 200 Hz if a first order filter is used or about 400 Hz for a second order Butterworth filter and about 600 Hz for a third order filter. Thus, the cut-off frequency can be computed as a function of the order of the filter using Equation 2:

$$\omega_C = 0.01nF_S \qquad (3)$$

In order to test crop attacks, the combined signal is cropped at different rates. It is observed that the watermark signal is retrievable only if the rate is very

close to 1 which is almost no cropping. With these results it is hard to say that the system is robust to crop attacks.

It is initially selected 4 digits to represent the DCT coefficients of the original and the watermark signals but this selection is changible to a lower resolution to match the low bandwidth of the transmission channel. However, each choice impacts the performance of the algorithm differently. If one or two digits of the DCT coefficients are used, like 0.0 or 0.9 for the previous example when the watermark signal DCT coefficient was 0.9876543, then the watermark signal is not retrievable where as three or more digits are used, i.e 0.98 or 0.987, the watermark signal is detectable by simulaton tool; however, it is more susceptible to channel constraints.

In the presence of white Gaussian noise (AWGN) which is simulated in Matlab, the watermark signal is retrievable at SNR greater than 135 dB for the case of the three digit DCT coefficients usage. It is detectable at 175 dB SNR for 4 digit representation of the DCT coefficients and at 210 dB SNR for the 5 digit representation. Here we may conclude that each digit corresponds to requiring an increase in the SNR of approximately 35 dB SNR. This response can be explained by looking at the number of digits needed to represent the DCT coefficient as the resolution of the signal. Higher resolutions are more susceptible to channel errors and, hence, require a higher SNR for error-free signal reconstruction.

When music is used as the message and the watermark signal, approximately 5 dB lower SNR gives successful watermark retrieval, as shown in Table 4. It means that speech is more fragile to noise than music.

Table 4. The precision of the DCT coefficient as a function of the SNR and the type of audio signal.

Sample SNR values at which the watermark signal is retrievable in the presence of an AWGN		
Digits used for DCT coefficents	SNR (speech)	SNR (music)
3 (i.e., 3.12)	135 dB	129 dB
4 (i.e., 3.123)	175 dB	170 dB
5 (i.e., 3.1234)	210 dB	211 dB

5. Conclusions and Future Work

In this work the main concern was to find a new solution to the problem of secure communications using digital audio watermarking. The novel approach presented here quantizes the DCT coefficients of the watermark signal that is to be hidden and embeds them into the DCT coefficients of the message, or host, signal. The results show that although the proposed algorithm has very good theoretical performance but since being susceptible to noise may not be easily implementable for real world application. However as future work, another approach based on attenuating, and then mixing, the DCT coefficients of the watermark and the original signal may be applied. The embedding algorithm will sum the host signal DCT coefficients with the watermark signal DCT coefficients. With this method one may construct a non-blind detection algorithm which uses a host signal at the receiver to recover the watermark signal but which will require more bandwidth, or one may construct a blind algorithm which does not use host signal, for retrieval of the watermark signal. Although the latter technique will be more complicated to implement, both approaches would be less susceptible to noise and likely to be more robust than our proposed algorithm.

6. References

[1] Cox I. J., Miller M., Bloom J., Friedrich J., and Kalker T., Digital Watermarking and Stenography, Morgan Kaufmann Publishers, 2008.

[2] Alsalami M. A. T., and Al-Akaidi M. M., "Digital Audio Watermarking: Survey," Proceedings of the 17th European Simulation Multiconference, 2003.

[3] Kim H. J., Choi, Y. H., Seok, J. W., and Hong, J. W., "Audio Watermarking Techniques," Intelligent Water Marking Techniques, Vol.7, No:1, pp. 185-218, 2004.

[4] Lee S.J., Jung S.H., "A survey of watermarking techniques applied to multimedia", Proceedings of IEEE Symposium on Industrial electronics, 2001.

[5] Xu C., Wu J., Sun Q., Xin K., 'Applications of Basamakal Watermarking Technology in Audio Signals', JAES Volume 47, Issue 10 pp. 805-812, 1999.

[6] Chenga Q., Wangb Y., Huangc T.S., 'Performance Analysis and Error Exponents of Asymmetric Watermarking Systems, Elsevier Signal Processing Volume 84, Issue 8, Pages 1429–1445, 2004.

[7] Wang X.Y., Zhao H., 'A Novel Synchronization Invariant Audio Watermarking Scheme Based on DWT and DCT', IEEE Transactions on Signal Processing, Vol. 54, No. 12, 2006.

[8] Bhat V.K., Sengupta I., Das A., 'Audio Watermarking Based on Mean Quantization in Cepstrum Domain', IEEE 2008.

[9] Ramalingam A., Krishnan S., 'Gaussian Mixture Modeling of Short-Time Fourier Transform Features for Audio Fingerprinting', IEEE Transactions On Information Forensics And Security, Vol. 1, No. 4, 2006.

[10] Megías D., Serra-Ruiz J., Fallahpour M., 'Efficient Self-Synchronised Blind Audio Watermarking System Based on Time Domain and FFT Amplitude Modification', Elsevier Signal Processing Volume 90, Issue 12, Pages 3078–3092, 2010.

Note:

Erol DUYMAZ was born in İzmir and received B.S. degree from 9th September University/ İZMİR in 2002. He joined Turkish Air Force (TurAF) in 2004 and worked in 1st ASMC/ Eskişehir until 2008 then he was assigned to Turkish Air Force Academy (TurAFA) ASTIN / İstanbul for M.S. degree and in 2009 for joint programme he attended ODU VA/USA and graduated from that university. He is still a PhD student in ASTIN TurAFA.

Aydın AKAN was born in Bursa and received B.S. degree from Uludağ University/ BURSA in 1988, M.S. degree from İstanbul Technical University in 1991 and PhD degree from University of Pittsburg PA/USA in 1996. He joined İstanbul University in 1996 and is still professor in Electrical and Electronics Engineering Department of that university.