Robust Audio Watermarking Using Multiwavelet Transform and Genetic Algorithm

PRAYOTH KUMSAWAT¹, KITTI ATTAKITMONGCOL² AND ARTHIT SRIKAEW² ¹School of Telecommunication Engineering, ²School of Electrical Engineering Institute of Engineering, Suranaree University of Technology 111 University Avenue, Muang District, Nakhon Ratchasima, 30000, Thailand E-mail: {prayoth, kitti, ra}@sut.ac.th

Abstract:- In this paper, we propose a new approach for optimization in digital audio watermarking using artificial intelligent technique. The watermarks are embedded into the low frequency coefficients in discrete multiwavelet transform domain. The embedding technique is based on quantization process which does not require the original audio signal in the watermark extraction. We have developed an optimization technique using the genetic algorithm to search for optimal quantization step in order to improve both quality of watermarked audio and robustness of the watermark. In addition, we analyze the performance of the proposed algorithm in terms of signal-to-noise ratio, normalized correlation and bit error rate. The experimental results show that our proposed method can improve the quality of the watermarked audio signal and give more robustness of the watermark as compared to previous works.

Key-Words: - Audio watermarking, Multiwavelet, Genetic algorithm, Artificial intelligence, Optimization

1 Introduction

Digital watermarking is one of the most popular approaches considered as a tool for providing the copyright protection of digital contents. This technique is based on direct embedding of additional information data into the digital contents. Ideally, there must be no perceptible difference between the watermarked and original digital contents, and the watermark should be easily extractable, reliable and robust against data compression or any signal manipulations [1]. The main requirements of digital watermarking are invisibility, robustness and data capacity. These requirements are mutually conflicting, and thus, in the design of a watermarking system, the trade off has to be made.

According to the International Federation of the Phonographic Industry (IFPI) audio [2], have watermarking should the following specifications: 1) Audio watermarking should not degrade perception of original signal. 2) Signal to noise ratio (SNR) should be greater than 20 dB and there should be more than 20 bits per-second (bps) data payload for watermark. 3) Watermark should be able to resist most common audio processing operations and attacks. 4) Watermark should be able to prevent unauthorized detection, removal and embedding, unless the quality of audio becomes very poor.

In general, digital audio watermarking can be performed in time domain and transform domain, where the properties of the underlying domain can be exploited. Currently, watermarking techniques based on transform domain are more popular than those based on time domain since they provide higher audio quality and much more robust watermark.

Seok and Hong [3] introduced direct sequence spread spectrum audio watermarking based on the discrete Fourier transform (DFT). The strength of the embedded watermark signal depends on the human perceptual characteristics of the audio signal. The detection procedure does not require access to the original audio signal to detect the watermark.

For many years, multi-scale representations and multiresolution analysis have proved useful in many signal processing applications. Wavelet analysis is an example to generate such representation [4, 5, 6]. In [6], Wang *et al.* proposed a digital audio watermarking algorithm based on the discrete wavelet transform (DWT). The watermark information is embedded in audio low-middle frequency coefficients in wavelet transform domain. A scheme of watermark detection is presented by using linear predictive coding, and it does not use the original signal during watermark extracting process. In [7], Chen and Wornell proposed a class of embedding methods called quantization index modulation (QIM) that achieves probably good rate-distortionrobustness performance. Wu et al. [8] proposed a self-synchronization algorithm for audio watermarking using QIM technique. They embed the synchronization codes with hidden informative the hidden data has selfdata SO that synchronization ability. Synchronization codes and informative bits are embedded into low-frequency subband in DWT domain. Their simulations suggest that the quantization step S (embedding strength) greatly depends on types and magnitudes of the original audio signals. It is not the best choice to use a fixed S. In [9], Kim and Bae proposed a robust algorithm to estimate the modified quantization step size with an optimal search interval. The equation to determine the optimal search interval is derived analytically, which can satisfy both detection performance and computational complexity. The authors conclude that the derived optimal search interval provides the accurate estimation of the modified quantization step size under amplitude modification attack.

In recent years, some multiwavelet-based digital watermarking algorithms have been proposed. Kwon and Tewfik [10] proposed an adaptive image watermarking scheme in the discrete multiwavelet transform (DMT) domain using successive subband quantization and a perceptual modeling. The watermark is Gaussian random sequence with unit variance and the original image is needed for watermark detection. Kumsawat et al. [11] proposed an image watermarking algorithm using the DMT and genetic algorithm is applied to search for optimal watermarking parameters to improve the quality of the watermarked image and the robustness of the watermark. Ghouti and Bouridane [12] proposed a novel audio fingerprinting framework for robust perceptual hashing of audio content using balanced multiwavelets. The extracted hash values are used for identifying, searching, and retrieving audio content from large audio databases. In [13], Kumsawat et al. proposed a multiwavelet-based audio watermarking scheme by utilizing the audio statistics characteristics and QIM technique. The watermarks are embedded into the low frequency coefficients in discrete multiwavelet transform domain to achieve robust performance against common signal processing procedures and noise corruptions.

Improvements in performance of digital audio watermarking schemes can be obtained by exploiting the characteristics of the human auditory system (HAS) in watermarking process. It is possible to embed perceptually inaudible watermarks with more energy in an audio, which makes watermark more robust [14]. Ding *et al.* [15] proposed an audio watermarking scheme based on wavelet packet and psychoacoustic model. The masking effects of the human auditory system are calculated in each subband by wavelet packet decomposition. The embedding strength is controlled by the masking threshold. Thus, the watermarking scheme has good secrecy and high robustness.

Another way to improve the performance of watermarking schemes is to make use of artificial intelligent (AI) techniques. The watermarking system can be viewed as an optimization problem. Therefore, it can be solved by Genetic Algorithm (GA), support vector machine (SVM), adaptive tabu search (ATS) or neural network (NN) [16]. There has been little research in application of GA to digital audio watermarking problems. Huang and Wu [17] proposed a watermarking method based on the discrete cosine transform (DCT) and Genetic Algorithm. They embed the watermark with visually recognizable patterns into the image by selectively modifying the middle-frequency parts of the image. The Genetic Algorithm is applied to search for the locations to embed the watermark in the DCT coefficient block such that the quality of the watermarked image is optimized. Sedghi et al. [18] proposed a novel approach in spread spectrum watermarking based on Genetic Algorithm for recovering Pseudo-noise (PN) sequence and detecting location of watermark signal without any information from the transmitter. Ketchem and Vongpradhip [19] presented audio watermarking technique using multiple image-based watermark scheme based on Genetic Algorithm in the DWT domain. They make use of Genetic Algorithm to find the optimum frequency bands for watermark embedding which can simultaneously improve robustness and audio quality of the watermarked audio. In [20], Srivingyong and Attakitmongcol proposed a robust audio watermarking method based on the DWT and the adaptive tabu search. Adaptive tabu search is the artificial intelligent searching technique designed for the solution of optimization problems. ATS is applied to search for optimal intensity of watermark such that the watermarked audio quality is optimized. Wang et al. [21] proposed a support vector machines-based audio watermarking scheme in wavelet domain. This algorithm embeds the template information and watermark signal into the original audio by adaptive quantization according to the local audio correlation and human auditory masking.

In this paper, we propose an audio watermarking method based on the discrete multiwavelet transform for the application of copyright protection. In our algorithm, the watermark is embedded into the multiwavelet transform coefficients using quantization index modulation technique. The watermark can be not only detected but also extracted to verify the owner. We apply the GA to search for optimal watermarking parameters in order to achieve optimum performance. Finally, we have compared the experimental results before and after optimization using GA with the results of previous works.

This paper is organized as follows: In Subsections 2.1 and 2.2, the preliminaries of multiwavelets and GA are introduced, respectively. Watermarking in the DMT domain with GA optimization is described in Section 3. In Section 4, the experimental results are shown. The conclusions of our study can be found in Section 5.

2 Preliminaries

2.1 Multiwavelet Transform

In recent years, multiwavelet transformation has gained a lot of attention in signal processing applications. The main motivation of using multiwavelet is that it is possible to construct multiwavelets that simultaneously possess desirable properties such as orthogonality, symmetry and compact support with a given approximation order [22, 23]. These properties are not possible in any scalar wavelet (wavelet based on one scaling function). One of the well-known multiwavelets was constructed by Donovan, Geronimo, Hardin, and Massopust (DGHM) [24]. DGHM multiwavelets simultaneously possess orthogonality, compact support, an approximation order of 2 and symmetry. A brief overview of the multiwavelet transform is described next.

Let Φ denotes a compactly supported orthogonal scaling vector $\Phi = (\phi^1, \phi^2, ..., \phi^r)^T$ where *r* is the number of scalar scaling functions. Then $\Phi(t)$ satisfies a two-scale dilation equation of the form

$$\Phi(t) = \sqrt{2} \sum_{n} h(n) \Phi(2t - n) \tag{1}$$

for some finite sequence h of $r \times r$ matrices. Furthermore, the integer shifts of the components of Φ form an orthonormal system, that is

$$\langle \phi^{l}(\cdot - n), \phi^{l'}(\cdot - n') \rangle = \delta_{l,l'} \delta_{n,n'}.$$
 (2)

Let V_0 denote the closed span of $\{\phi^l(\cdot - n) \mid n \in \mathbb{Z}, l = 1, 2, ..., r\}$ and define $V_j = \{f(\frac{\cdot}{2^j}) \mid f \in V_0\}$. Then $(V_j)_{j \in \mathbb{Z}}$ is a

multiresolution analysis of $L^2(R)$. Note that the decreasing convention $V_{i+1} \subset V_i$ is chosen.

Let W_j denotes the orthogonal complement of V_j in V_{j-1} . Then there exists an orthogonal multiwavelet $\Psi = (\psi^1, \psi^2, ..., \psi^r)^T$ such that $\{\psi^l(\cdot - n) | l = 1, 2, ..., r$ and $n \in Z\}$ form an orthonormal basis of W_0 . Since $W_0 \subset V_{-1}$, there exists a sequence g of $r \times r$ matrices such that

$$\Psi(t) = \sqrt{2} \sum_{n} g(n) \Phi(2t - n).$$
(3)

Let $f \in V_0$, then f can be written as a linear combination of the basis in V_0 :

$$f(t) = \sum_{n} c_0(k)^T \Phi(t - k)$$
 (4)

for some sequence $c_0 \in l_2(Z)^r$. Since $V_0 = V_1 \oplus W_1$, *f* can also be expressed as

$$f(t) = \frac{1}{\sqrt{2}} \sum_{k \in \mathbb{Z}} c_1(k)^T \Phi(\frac{t}{2} - k) + \frac{1}{\sqrt{2}} \sum_{k \in \mathbb{Z}} d_1(k)^T \psi(\frac{t}{2} - k).$$
(5)

The coefficients c_1 and d_1 are related to c_0 via the following decomposition and reconstruction algorithm:

$$c_1(k) = \sum_n h(n)c_0(2k+n)$$
(6)

$$d_1(k) = \sum_{n} g(n)c_0(2k+n)$$
(7)

$$c_{0}(k) = \sum_{n} h(k - 2n)^{T} c_{1}(n) + \sum_{n} g(k - 2n)^{T} d_{1}(n).$$
(8)

Unlike scalar wavelet, even though the multiwavelet is designed to have approximation order p, the filter bank associated with the multiwavelet basis does not inherit this property. Thus, in applications, one must associate a given discrete signal into a sequence of length -r vectors without losing some certain properties of the underlying multiwavelet. Such a process is referred to as prefiltering. The block diagram of a multiwavelet with prefilter Q(z) and postfilter P(z)

is shown in Fig. 1 where c_1 is the approximation subband which mainly represents the low frequency component of the audio signal, and d_1 is the detail subband which mainly represents the high frequency component of the audio signal. H(z) and G(z) are the z transform of h(n) and g(n), respectively. Two audio subbands are obtained from each level of decomposition; one detail subband and one approximation subband. For the next level of decomposition, the multiwavelet transform is applied to the approximation subband of the previous decomposition level. Thus, n levels of decomposition result in n+1 subbands at the analysis filter bank.



Fig.1 Multiwavelet filter bank

2.2 Genetic Algorithm

Genetic algorithm (GA) is one of the most widely used artificial intelligent techniques belonging to the area of evolutionary computation. Genetic algorithm based on the mechanisms of natural selection and genetics, has been developed since 1975 [25] and has been applied to a variety of optimization and search problems [11, 17, 18, 19, 23]. GA has been proven to be very efficient and stable in searching for global optimum solutions.

Usually, a simple GA is mainly composed of three operations: selection, genetic operation and replacement. A brief summary for implementing GA can be summarized as follows:

Defining the solution representation of the system is the first task of applying GA. GA uses a population, which is composed of a group of chromosomes, to represent the solutions of the system. The solution in the problem domain can then be encoded into the chromosome in the GA domain and vice versa. Initially, a population is randomly generated. The fitness function then uses objective values from objective function to evaluate the fitness of each chromosome. The fitter chromosome has the greater chance to survive during the evolution process.

The objective function is problem-specific; its objective value can represent the system performance index (e.g. an error). Next, a particular group of chromosomes is chosen from the population to be parents. The offspring is then generated from these parents by using genetic operations, which normally are crossover and mutation. Similar to their parents, the fitness of the offspring is evaluated and used in replacement processes in order to replace the chromosomes in the current population by the selected offspring.

The GA cycle is then repeated until a desired termination criterion is satisfied, for example, the maximum number of generations is reached or the objective value is below the threshold. There are various techniques in designing GA that we have to take into account. These include encoding schemes, fitness evaluation, parent selection, genetic operations and replacement strategies.

3 Proposed Method

In this section, we first give a brief overview of the watermark embedding and watermark extracting algorithms in the DMT domain based on the concept of the quantization index modulation technique. We then describe our proposed optimization technique in watermarking scheme using genetic algorithm.

3.1 Watermark Embedding Algorithm

The watermark embedding algorithm is described as follows:

1. Generate a seed by mapping a signature or text through a one-way deterministic function. The seed is used as the secret key for watermarking.

2. To increase security, perform a pseudorandom permutation in order to disperse the spatial relationship of the binary watermark pattern. Therefore, it would be difficult for a pirate to detect or remove the watermark. We use W and \overline{W} to denote the original watermark image and the permuted watermark image, respectively. The relationship between W and \overline{W} can be expressed as $\overline{W}(i, j) = W(i', j')$, where (i', j') is permuted to the pixel position (i, j) in a secret order using the secret key. Since the audio signal is onedimensional, we should transform the permuted watermark image into the one-dimensional sequence in order to embed it in the audio signal. Then, the \overline{W} is transformed and mapped into a binary antipodal sequence $\hat{W} = \{\hat{w}_i\}$ for $i = 1, 2, ..., N_w$, where N_w is the length of watermark and $\hat{w}_i \in \{+1, -1\}$.

3. Transform the original audio signal into fivelevel decomposition using the DMT. Since the approximation coefficients are supposed to be relatively stable and less sensitive to slight changes of the audio signal, they are ideal embedding area. In order to achieve a balance between robustness and fidelity, the coefficients at coarsest approximation subband are selected for watermark embedding based-on artificial intelligent technique. Furthermore, the coefficients in high-frequency subband are not used for watermark embedding because of their low signal energy in this frequency band.

4. Select the significant coefficients in the DMT domain which is the first N_w largest coefficients at coarsest approximation subband to embed the watermark bits. The position of significant coefficients will be sent to the receiver as the side information. To increase the watermarking security, we order the N_w largest coefficients in a pseudorandom manner. The random numbers can be generated using the same secret key in step (1).

5. For watermark embedding, the sequence $\{\hat{w}_i\}$ is embedded into the selected coefficients by quantization index modulation technique. The quantization function is given as follows:

$$c'_{i} = \begin{cases} \lfloor c_{i} / S \rfloor \cdot S + 3S / 4 & if \ \hat{w}_{i} = +1 \\ \lfloor c_{i} / S \rfloor \cdot S + S / 4 & if \ \hat{w}_{i} = -1 \end{cases}$$
(9)

, where $\lfloor x \rfloor$ rounds to the greatest integer smaller than x, $\{c_i\}$ and $\{c'_i\}$ are the DMT coefficients of the original audio data and the corresponding watermarked audio data respectively, and S is quantization step. A large S makes the watermark robust, but it will destroy the original quality of the audio. Thus, the value of S should be as large as possible under the constraint of imperceptibility. In order to improve both quality of watermarked audio and robustness of the watermark, this work employs the Genetic Algorithm to search for the optimal quantization step. This quantization step is varied to achieve the most suitable watermarked audio signal for each given audio signal. The details of GA optimization process will be described in details in Section 3.3.

6. Perform inverse DMT to obtain the watermarked audio signal. The overall watermark embedding process is shown in Fig. 2.



Fig. 2 Watermark embedding process

3.2 Watermark Extracting Algorithm

The watermark extracting algorithm is outlined as follows:

1. Transform the watermarked audio signal into five-level decomposition using the DMT to obtain detail coefficients and approximation coefficients. Then, we choose the first N_w largest coefficients in the coarsest approximation subband from position in the side information. We further order the N_w largest coefficients in a pseudorandom manner using the secret key.

2. Let \tilde{c}_i denote the N_w largest coefficients of the coarsest approximation subband. The embedded watermark can be extracted from \tilde{c}_i by using the following rule:

$$\widetilde{w}_{i}^{*} = \begin{cases} +1 & if \quad \widetilde{c}_{i} - \lfloor \widetilde{c}_{i} / S \rfloor \cdot S \ge S / 2\\ -1 & if \quad \widetilde{c}_{i} - \lfloor \widetilde{c}_{i} / S \rfloor \cdot S < S / 2 \end{cases}$$
(10)

3. Inverse the permutation of \tilde{W}^* where $\tilde{W}^* = \{\tilde{w}_i^*, i = 1, 2, ...N_w\}$ to obtain the extracted watermark \tilde{W} . In our proposed method, the extracted watermark is a visually recognizable image. After extracting the watermark, we used normalized correlation coefficients to quantify the correlation between the original watermark and the extracted one. A normalized correlation (*NC*) between *W* and \tilde{W} is defined as:

$$NC(W, \widetilde{W}) = \frac{\sum_{i=1}^{N_w} w_i \widetilde{w}_i}{\sqrt{\sum_{i=1}^{N_w} w_i^2 \sum_i \widetilde{w}_i^2}}$$
(11)

, where W and \tilde{W} denote an original watermark and extracted one, respectively and $\tilde{W} = \{\tilde{w}_i\}$ for $i = 1, 2, ..., N_w$. The watermark extracting process is shown in Fig. 3.



Fig. 3 Watermark extracting process

3.3 Improving Performance using Genetic Algorithms

In the design of digital audio watermarking system, there are three goals that are always conflicted. These goals are imperceptibility, robustness and data capacity. In order to minimize such conflicts, this work employs the genetic algorithm to search for an optimal watermarking parameters. This For the optimization process, GA is applied in the watermark embedding and the watermark extracting processes to search for quantization step (S). The objective function of searching process is computed by using factors that relate to both robustness and imperceptibility of a watermark. A high quality output audio and robust watermark can then be achieved. The diagram of our proposed algorithm of applying GA is shown in Fig. 4 and details of GA are described as follows:



Fig. 4 Optimization diagram for digital audio watermarking using genetic algorithm

Chromosome Encoding: Chromosomes in GA represent desired parameter to be searched. Number of chromosomes used in this work is 20. The encoding scheme is binary string with 32 bit resolutions for each chromosome. Hence, the parameter S is represented by chromosome with length of 32 bits.

Objective Function Evaluation: The most critical step in the GA optimization process is the definition of a reliable objective function. In this paper, the objective function of GA uses both normalized correlation (*NC*) and difference (*DIF*) between desired signal-to-noise ratio (*SNR*) and obtained *SNR* from each iteration as performance indexes. *DIF* is an imperceptibility measure, while *NC* is a robustness measure.

According to the International Federation of the Phonographic Industry, the *SNR* of watermarked

audio signal should be greater than 20 dB. Therefore, the value of desired *SNR* has been assigned to 24 dB in all experiments. During GA-based optimization processes, three attacks are chosen to evaluate the imperceptibility and robustness of the embedded watermark. They are MP3 compression at 64 kbps, Gaussian noise addition, and re-quantization. Details of these attacks will be thoroughly described in Section 4.3. After obtaining the *SNR* in the watermarked audio, the *DIF* value and the average of the three normalized correlations (NC_{ave}) after attacking, we are ready to start the objective function evaluation. An illustrative diagram is shown in Fig. 4. The objective function f_{obj} can be evaluated as follow:

$$f_{obj} = \delta_{DIF} \times DIF + \delta_{NC} \times NC_{ave}$$
(12)

, where δ_{DIF} and δ_{NC} are weighting factors of *DIF* and *NC*_{ave}, respectively. These weighting factors represent the significance of each index used in GA searching process. If both indexes are equally significant, the values of these factors will be 0.5 each where the relationship $\delta_{DIF} + \delta_{NC} = 1.0$ must always hold. In this work, the weighting factors δ_{DIF} and δ_{NC} are equally set to 0.5.

In order to gain the optimal performance of the quantization-based audio watermarking system, f_{obj} should be optimized at GA processes. By using objective function f_{obj} above, the parameter *S* can be optimally searched to achieve the best of both output audio quality and watermark robustness.

Selection, Genetic Operation, and Replacement:

After evaluating fitness value of each chromosome based on the proposed objective function, chromosomes will be selected to produce offsprings by crossover and mutation operations. In this work, a ranking selection is chosen for selection mechanism. The crossover is uniform, with probability of 0.7. Mutation is standard, with probability of 0.05. The chromosomes are then partially replaced by the best chromosome for each generation.

The GA will be iteratively performed on an input audio signal until a desired termination is satisfied. In this work, the maximum number of generations is set to 30 as our stopping criterion. Then the chromosome (the solution) with the best fitness value, i.e., the quantization step S, is determined.

4 Experimental Results and Discussions

In order to demonstrate the performance of the proposed algorithm, some numerical experiments are carried out to measure the audio quality of the watermarked audio and evaluate the robustness of the watermark under typical attacks.

A set of ten audio signals have been used as host signals, representing five general classes of music: classical, country, jazz, rock, and pop. This delineation has been chosen because each class has different spectral properties. Each audio signal has duration of 30 seconds in the WAV format and is mono, 16 bits/sample, with sampling rate of 44.1 kHz. A binary logo "SIP SUT" of size 32×32 pixels (N_w = 1,024) is used as the visually recognizable watermark. Consequently, the total watermark data rate is 34.14 bps which satisfies the IFPI requirement described in Section 1. Figures 5(a) and (b) show the original watermark and permuted watermark, respectively.

We use *SNR* (Signal-to-noise ratio), *NC* (Normalized correlation) and *BER* (Bit error rate) to analyze the performance of the proposed algorithm. The *BER* and *SNR* are defined as:

$$BER = \frac{Number of \ error \ bits}{Number \ of \ total \ bits} \times 100\%$$
(13)

$$SNR = 10\log_{10} \left(\frac{\sum_{i} (f_{i}^{2})}{\sum_{i} (f_{i} - f_{i}^{\prime})^{2}} \right)$$
(14)

, where f_i and f'_i denote the original and modified audio, respectively.



Fig. 5 (a) Original watermark and (b) permuted watermark

4.1 Results of Genetic Algorithm Optimization

Figure 6 shows the convergence of GA optimization at 30 generations of the Pop2 audio signal. It is obvious that as the number of generation increases, the improvement of audio quality (SNR) gradually approaches to a saturation value. The resulting parameters, which are quantization steps *S* from GA optimization of 10 test audios, are shown in the Table 1. These parameters are optimally varied to achieve the most

desirable ones for original audios with different characteristics.

4.2 Imperceptibility Test Results

The watermarked audio quality is examined by watermarking the original audio signals with the resulting parameters from GA. Then, the *SNR* test is conducted, which serves as an objective measurement of audio signal quality. The *SNR* is measured by comparing the watermarked signal with the original one.



Fig. 6 *SNR* , *DIF* , NC_{ave} and f_{obj} from GA optimization process

 Table 1 Parameter S from GA process of each audio signal

Host signal	S
Classical1	0.3211
Classical2	0.3067
Country1	0.3519
Country2	0.3033
Jazz1	0.3067
Jazz2	0.3067
Rock1	0.3219
Rock2	0.3165
Pop1	0.3067
Pop2	0.3063

Figure 7 shows the original Classical1 audio signal waveform and the corresponding watermarked audio signal waveform. Note that *SNR* as high as 27.39 dB for watermarked audio signal. However, there is no obvious difference between original signal and watermarked signal by using informal listening test, and from Fig. 7(a) and (b). It demonstrates that the proposed algorithm has perfect insensibility in the sense of hearing.

The results of watermarked audio quality are shown in Table 2. The results obtained from our proposed method which is called With-GA (After optimization) are compared with the method without using Genetic Algorithm which is referred to as Without-GA (Before optimization). In Without-GA method, the quantization step is fixed at 0.4. We can see that the proposed method can improve the *SNR* of the watermarked audio about 2 dB.



Fig. 7 (a) Original audio signal (Classical1), (b) watermarked audio signal

Table 2	Signal-to-	noise rati	o of	watermarked	audio
signals					

	SNR (dB)				
Host signals	Without-GA	With-GA			
Classical1	25.31	27.39			
Classical2	22.23	23.84			
Country1	28.60	30.06			
Country2	23.06	25.42			
Jazz1	23.03	25.05			
Jazz2	26.09	28.29			
Rock1	26.73	28.24			
Rock2	27.99	29.83			
Pop1	21.82	23.55			
Pop2	22.12	24.94			
Average	24.69	26.66			

4.3 Robustness Test Results

We first tested the robustness of the proposed algorithm to 10 audio samples under no attacks. If the *BER* of the recovered watermark sequence is 0, it means that the embedded bit can be recovered exactly. The effects of the following six types of attacks are then investigated.

1. Re-sampling: The audio signal is first down-sampled at 22.05 kHz, and then up-sampled at 44.1 kHz.

2. Re-quantization: The 16-bit watermarked audio signals have been re-quantized down to 8 bits/sample and back to 16 bits/sample.

3. Low-pass filtering: Low-pass filtering using a second order Butterworth filter with cut-off frequency of 6 kHz, 12 dB/octave roll-off, is performed to the watermarked audio signals.

4. Addition of noise: White Gaussian noise with 1% of the power of the audio signal is added.

5. Cropping: Two thousand samples of each testing signal are cropped out at 5 random positions.

6. Low bit-rate codec: The robustness against the low-rate codec was tested by using MPEG 1 Layer III compression (MP3) with compression rates of 56, 64, 96, and 128 kbps.

Detection results for the various attacks described above are shown in Table 3 which displayed the *NC* and *BER* from watermark

extraction. The experimental results given in Table 3 show that the watermark is not affected by resampling, re-quantization, additive noise, and MP3 compression at 64, 96, and 128 kbps. This indicated that the watermark is very robust to these attacks.

For low-pass filtering, cropping and MP3 compression at 56 kbps attacks, the *BER* values of the recovered watermark sequence are 8.0133%, 5.2636% and 0.0977% for the Without-GA method and 7.0313%, 5.2539% and 0.0977% for the With-GA method, respectively. Although a lot of loss occurred in the audio signal, the bit error rates are still acceptable. The results show that our proposed method yields better results than the method without GA. Because GA search guarantees the global optimum solution, the proposed method can thus improve the quality of the watermarked audio and give almost the same robustness of the watermark.

	Without-GA		With-GA	
Attack type	NC	BER	NC	BER
		(%)		(%)
Attack free	1.0000	0.0000	1.0000	0.0000
Re-sampling	1.0000	0.0000	1.0000	0.0000
Re-quantization	1.0000	0.0000	1.0000	0.0000
Low-pass	0.9292	8.0133	0.9369	7.0313
filtering				
Additive noise	1.0000	0.0000	1.0000	0.0000
Cropping	0.9031	5.2636	0.9529	5.2539
MP3-56kpbs	0.9992	0.0977	0.9992	0.0977
MP3-64kpbs	1.0000	0.0000	1.0000	0.0000
MP3-96kpbs	1.0000	0.0000	1.0000	0.0000
MP3-128kpbs	1.0000	0.0000	1.0000	0.0000

 Table 3 Robustness comparison of our algorithm

Finally, results obtained from our proposed method which is called With-GA are compared in finer details with the method based on wavelet transform and ATS in [20]. In order to compare robustness between the two techniques in a fair manner, parameters for each scheme should be adjusted so that watermarked audio signals of approximately close imperceptibility are produced. In these experiments, the *SNR* of watermarked audio in each scheme has been set to 24 dB. According to the experimental results, the value of the embedding capacity has been assigned to 34.14 bps in all experiments. The comparison results are listed in Table 4.

Table 4 shows test results of Rock1 audio signal with no attack, re-sampling, re-quantization, low-pass filtering, addition of noise, cropping and MPEG 1 Layer III compression with compression rates of 56, 64, 96 kbps and 128 kbps, respectively. The *BER* of watermark image and the *SNR* of digital audio signal are also displayed.

	With-GA		[20]	
Attack type	BER	SNR	BER	SNR
	(%)	(dB)	(%)	(dB)
Attack free	0.0000	24.1673	0.0000	24.1209
Re-sampling	0.0000	23.8597	0.0000	23.5533
Re-quantization	0.0000	20.5791	0.1953	20.0936
Low-pass	1.1719	10.3440	3.3203	10.3506
filtering				
Additive noise	0.0000	19.6423	2.6367	19.5607
Cropping	4.0039	10.5865	5.3711	10.8289
MP3-56kpbs	0.0000	18.7144	5.7617	18.5400
MP3-64kpbs	0.0000	18.8226	5.3711	18.5348
MP3-96kpbs	0.0000	19.6279	4.1016	19.5537
MP3-128kpbs	0.0000	19.7944	3.6133	19.5607

According to these results, the extracted watermark images from our proposed method have some distortion for low-pass filtering and cropping attacks but they are still visually recognizable. In addition, the bit error rates of the extracted watermarks using our proposed method are always lower than the ones using method in [20]. The results demonstrate that our proposed method yields significantly more robust watermark than the method in [20] does.

5 Conclusion

digital This paper proposes a audio watermarking algorithm in the multiwavelet transform domain. In order to make the watermarked signal inaudible, the watermark is embedded into low frequency part of the highest energy of audio signal by taking advantage of multi-resolution characteristic of multiwavelet transform. The watermark insertion and watermark extraction are based on the quantization index modulation technique and the watermark extraction algorithm does not need the original audio in the extraction process. We have developed an optimization technique using the genetic algorithm. In our optimization process, we use genetic algorithm searching for optimal parameter which is the quantization step. This parameter is optimally varied to achieve the most suitable for original audios with different characteristics. The testing results of the watermarked audio quality and watermark robustness with various watermark attacks show that our proposed method can improve the performance of the watermarking process such that the better watermarked audio quality and watermark robustness are achieved. Further research can be concentrated on the development of our proposed method by using the characteristics of the human auditory system.

Acknowledgement

This work was supported by a grant from Suranaree University of Technology, Nakhon Ratchasima, Thailand. The authors would like to thank the Thailand Research Fund and Commission on Higher Education for all great supports.

References:

- S. J. Lee and S. H. Jung, "A Survey of Watermarking Techniques Applied to Multimedia," Proc. IEEE Int. Symp. Industrial Electronics, Pusan, South Korea, vol. 1, pp. 272-277, June 2001.
- [2] S. Katzenbeisser and F. A. P. Petitcolas, "Information Hiding Techniques for Steganography and Digital Watermarking," *Artech House*, Massachusetts, 2000.
- [3] J. W. Seok and J. W. Hong, "Audio Watermarking for Copyright Protection of Digital Audio Data," *IEE Electronics Letters*, vol. 37, pp. 60-61, 2001.
- [4] M. Shinohara, F. Motoyoshi, O. Uchida and S. Nakanishi, "Wavelet-Based Robust Digital Watermarking Considering Human Visual System," WSEAS Trans. on Signal Processing, vol. 3, pp. 185-190, Feb. 2007.
- [5] X. Wang, Y. O. Yang, and H. M. Gu, "A Remote Sensing Image Self-Adaptive Blind Watermarking Algorithm Based on Wavelet Transformation," Proc. 7th WSEAS Int. Conf. on Signal, Speech and Image Processing, Beijing, China, vol.1, pp. 76-82, Sept. 2007.
- [6] R. Wang, D. Xu, J. Chen and C. Du, "Digital Audio Watermarking Algorithm Based on Linear Predictive Coding in Wavelet Domain," IEEE Int. Conf. Signal Processing, Beijing, China, vol. 3, pp 2393-2396, August 2004.
- [7] B. Chen and G. Wornell, "Quantization Index Modulation: A Class of Provably Good Methods for Digital Watermarking and Information Embedding," *IEEE Trans. on Information Theory*, vol. 47 no. 4, pp. 1423-1443, May 2001.
- [8] S. Wu, J. Huang, D. Huang and Y. Q. Shi, "Efficiently Self-Synchronized Audio Watermarking for Assured Audio Data Transmission," *IEEE Trans. on Broadcasting*, vol. 51. no. 1, pp. 69–76, March 2005.
- [9] S. Kim and K. Bae, "Analysis of Optimal Search Interval for Estimation of Modified Quantization Step Size in Quantization-Based Audio Watermark Detection," *Lecture Notes in Computer Science*, 4139, pp. 604-615, 2006.
- [10] K. R. Kwon and A. H. Tewfik, "Adaptive Watermarking Using Successive Subband Quantization and Perceptual Model Based on Multiwavelet Transform," Proc. SPIE Int.

Prayoth Kumsawat, Kitti Attakitmongcol, Arthit Srikaew

Conf. Security and Watermarking of Multimedia Contents IV, San Jose, CA, USA., vol. 4675, pp. 334-348, April 2002.

- [11] P. Kumsawat, K. Attakitmongcol and A. Srikaew, "A New Approach for Optimization in Image Watermarking by Using Genetic Algorithms," *IEEE Trans. Signal Processing*, vol. 53, pp. 4707-4719, 2005.
- [12] L. Ghouti and A. Bouridane, "A Robust Perceptual Audio Hashing Using Balanced Multiwavelets," Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing, Toulouse, France, vol. 1, pp. 209-212, May 2006.
- [13] P. Kumsawat, K. Attakitmongcol and A. Srikaew, "Digital Audio Watermarking for Copyright Protection Based on Multiwavelet Transform," *Lecture Notes in Computer Science*, vol. 5376, pp. 155-164, 2008.
- [14] F. Hartung and M. Kutter, "Multimedia Watermarking Techniques," Proc. IEEE, vol. 87, no. 7, pp. 1079-1107, July 1999.
- [15] W. R. Ding, X. D. Wen and L. Qian, "Audio Watermarking Algorithm Based on Wavelet Packet and Psychoacoustic Model," Proc. Sixth Int. Conf. Parallel and Distributed Computing, Applications and Technologies, Dalian, China, vol. 1, pp. 812-814, Dec. 2005.
- [16] Y. Zhang, "Blind Watermark Algorithm Based on HVS and RBF Neural Network in DWT Domain," WSEAS Trans. on Computers, vol. 8, pp. 174-183, Jan. 2009.
- [17] C. H. Huang and J. L. Wu, "A Watermark Optimization Technique Based on Genetic Algorithms," Proc. SPIE Int. Conf. Visual Communications and Image Processing, San Jose, CA, USA., vol. 3971, pp. 516-523 Feb. 2000.
- [18] S. Sedghi, H. R. Mashhadi and M. Khademi, "Detecting Hidden Information from a Spread Spectrum Watermarked Signal by Genetic Algorithm," Proc. IEEE Conf. Evolutionary Computation, Vancouver, BC, Canada, vol. 1, pp. 173-178, July 2006.
- [19] M. Ketcham and S. Vongpradhip, "Genetic Algorithm Audio Watermarking Using Multiple Image-Based Watermarks," Proc. Int. Symp. Communications and Information Technologies, Sydney, Australia, vol. 1, pp. 1235-1240, Oct. 2007.
- [20] N. Sriyingyong and K. Attakitmongcol, "Wavelet-Based Audio Watermarking Using Adaptive Tabu Search," Proc. IEEE Int. Symp. Wireless Pervasive Computing, Phuket, Thailand, vol.1, pp. 1-5 Jan. 2006.
- [21] X. Y. Wang, P. P. Niu and H. Y. Yang, "A Robust, Digital-Audio Watermarking

Method," *IEEE Multimedia*, vol. 16, pp. 60-69, 2009.

- [22] K. Attakitmongcol, D. P. Hardin and D.M. Wilkes, "Multiwavelet Prefilters II: Optimal Orthogonal Prefilters," *IEEE Trans. Image Processing*, vol. 10, no. 10, pp. 1476-1487, 2001.
- [23] P. Kumsawat, K. Attakitmongcol and A. Srikaew, "An Optimal Robust Digital Image Watermarking Based on Genetic Algorithms in Multiwavelet Domain," WSEAS Trans. Signal Processing, Issue 1, vol. 5, pp. 42-51, Jan. 2009.
- [24] S. J. Geronimo, D. P. Hardin and P. R. Massopust, "Fractal Functions and Wavelet Expansions Based on Several Scaling Functions," *Approximation Theory Journal*, vol. 78, pp. 373-401, 1994.
- [25] J. H. Holland, "Adaptation in Natural and Artificial Systems," Ann Arbor, The University of Michigan Press, 1975.