Research Article

A Genetic Algorithm Optimization Technique for Multiwavelet-Based Digital Audio Watermarking

Prayoth Kumsawat

School of Telecommunication Engineering, Institute of Engineering, Suranaree University of Technology, 111 University Avenue, Muang District, Nakhon Ratchasima 30000, Thailand

Correspondence should be addressed to Prayoth Kumsawat, prayoth@sut.ac.th

We propose a new approach for optimization in digital audio watermarking using genetic algorithm. 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 four optimal quantization steps 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 the proposed scheme can achieve a good robustness against most of the attacks which were included in this study.

1. Introduction

During the past few years, digital multimedia technology and Internet networks have made great progress and they are now becoming increasingly important in daily life. Consequently, intellectual property protection is a pressing concern for content owners who are exhibiting digital representation of the photographs, music, video, and original artworks through the Internet.

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 (called watermark) 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 tradeoff has to be made.

According to the International Federation of the Phonographic Industry (IFPI) [2], audio watermarking should have 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.

Typical schemes for digital watermarking were based on transform-domain techniques with discrete cosine transform (DCT) [3, 4], discrete Fourier transform (DFT) [5, 6], and discrete wavelet transform (DWT) [7–12].

In recent years, some discrete multiwavelet transform-based digital watermarking algorithms have been proposed. Kwon and Tewfik [13] proposed an adaptive image watermarking scheme in the discrete multiwavelet transform

(DMT) domain using successive subband quantization and a perceptual modeling.

Kumsawat et al. [14] proposed an image watermarking algorithm using the DMT and genetic algorithm (GA) applied to search for optimal watermarking parameters to improve the quality of the watermarked image and the robustness of the watermark. Ghouti and Bouridane [15] proposed a novel audio fingerprinting framework for robust perceptual hashing of audio content using balanced multi-wavelets. In [16], Kumsawat et al. proposed a multiwavelet-based audio watermarking scheme by utilizing the audio statistics characteristics and quantization index modulation (QIM) technique.

Improvements in performance of digital audio watermarking schemes can be obtained by exploiting the characteristics of the human auditory system in watermarking process. It is possible to embed perceptually inaudible watermarks with more energy in an audio, which makes watermark more robust [12, 17].

Another way to improve the performance of watermarking schemes is to make use of artificial intelligence (AI) techniques. The watermarking system can be viewed as an optimization problem. Therefore, it can be solved by genetic algorithm [4, 14, 18], adaptive tabu search (ATS) [10], or support vector machine (SVM) [11]. There has been little research in application of GA to digital audio watermarking problems.

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 four 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 Sections 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 and discussions 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 [14, 19]. 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) [20]. DGHM multiwavelets simultaneously possess orthogonality, compact support, an approximation order of 2, and symmetry.
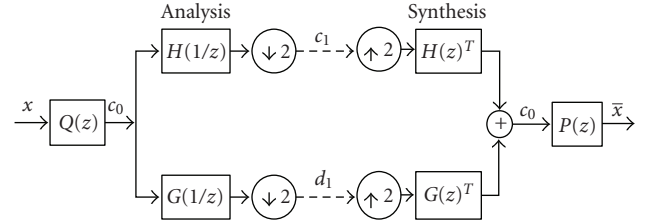

FIGURE 1: Multiwavelet filter bank.

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. Furthermore, since the multiwavelets have more than one scaling function, the dilation equation becomes a dilation with matrix coefficients. Thus, in applications, one must associate a given discrete signal into a sequence of length $-r$ vectors (where r is the number of scaling functions) without losing some certain properties of the underlying multiwavelet. Such a process is referred to as prefiltering or multiwavelet initialization.

The block diagram of a multiwavelet with prefilter $Q(z)$ and postfilter $P(z)$ is shown in Figure 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.

2.2. Genetic Algorithm. Genetic algorithm 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 [21] and has been applied to a variety of optimization and search problems [4, 14, 18, 22]. 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
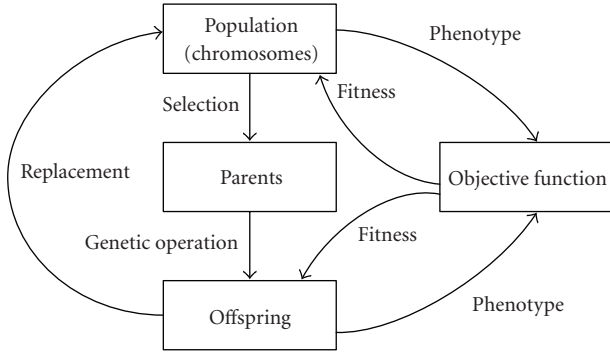
FIGURE 2: GA Cycle.

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. The GA cycle is shown in Figure 2 where the phenotype is the coding scheme used to represent the chromosomes.

3. The Proposed Watermarking Algorithm

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 (K) 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 (K). Since the audio signal is one-dimensional, 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 $\widehat{W} = \{\widehat{w}_i\}$ for $i = 1, 2, \ldots, N_w$, where N_w is the length of watermark and $\widehat{w}_i \in \{+1, -1\}$.

(3) Transform the original audio signal using the DMT. Five-level DMT is chosen because it was found in our experimental results that the higher levels do not yield significantly better results while they increase computational complexity. 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.

(5) Based on the statistical values of the selected coefficients, we have grouped them into 4 groups corresponding to their magnitudes. Each group will be quantizing with different quantization steps. To increase the watermarking security, we order the selected coefficients in a pseudorandom manner. The random numbers can be generated using the same secret key (K) in generating the watermark $\widehat{W}$.

(6) For watermark embedding, the sequence $\{\widehat{w}_i\}$ is embedded into the selected coefficients by quantization index modulation technique. The quantization function is given as follows [16]:

$$c_i' = \begin{cases} \left\lfloor \dfrac{c_i}{S_j} \right\rfloor \cdot S_j + \dfrac{3S_j}{4}, & \text{if } \widehat{w}_i = +1, \\[4mm] \left\lfloor \dfrac{c_i}{S_j} \right\rfloor \cdot S_j + \dfrac{S_j}{4}, & \text{if } \widehat{w}_i = -1, \end{cases} \tag{1}$$

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. The variable S_j, for $j = 1, 2, 3, 4$, denotes the quantization steps corresponding to 4 groups of the selected coefficients. The coarsest approximation subband of five-level multiwavelet decomposition and the 4 groups of the first N_w largest coefficients are shown in Figures 3(a) and 3(b), respectively.

A large quantization step makes the watermark robust, but it will destroy the original quality of the audio signal. Thus, the value of quantization step should be as large as possible under the constraint of imperceptibility.

(7) In order to optimize both quality of watermarked audio and robustness of the watermark, this work employs the genetic algorithm to search for 4 optimal quantization steps. These quantization steps are varied to achieve the most suitable watermarked audio signal for each given audio signal. The details of genetic algorithm optimization process will be described in details in Section 3.3.
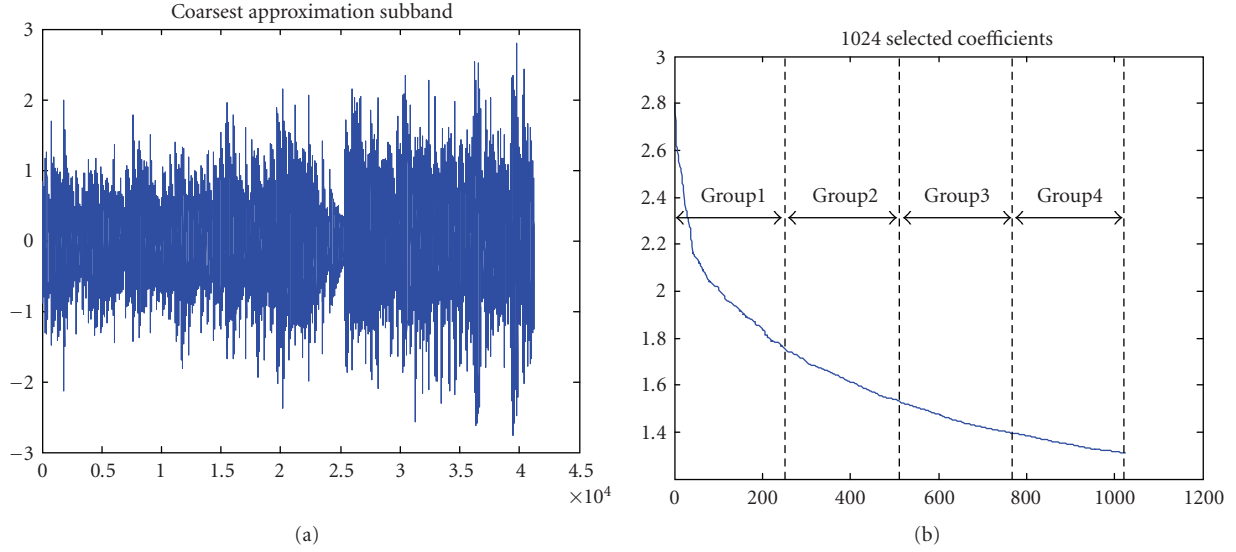
FIGURE 3: (a) The coarsest approximation subband of five-level multiwavelet decomposition and (b) the 4 groups of the first $N_w(1024)$ largest coefficients.

(8) Reorder the selected coefficients and inverse DMT to obtain the watermarked audio signal. The overall watermark embedding process is shown in Figure 4.

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 detailed coefficients and approximation coefficients. Then, we choose the first N_w largest coefficients in the coarsest approximation subband from position in the side information and group them into 4 groups. We further order them in a pseudorandom manner using the secret key (K).

(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:

$$\tilde{w}_i^* = \begin{cases} +1, & \text{if } \tilde{c}_i - \left\lfloor \dfrac{\tilde{c}_i}{S_j} \right\rfloor \cdot S_j \geq \dfrac{S_j}{2}, \\ \\ -1, & \text{if } \tilde{c}_i - \left\lfloor \dfrac{\tilde{c}_i}{S_j} \right\rfloor \cdot S_j \geq \dfrac{S_j}{2}. \end{cases} \tag{2}$$

(3) Inverse the permutation of $\widetilde{W}^*$ where $\widetilde{W}^* = \{\tilde{w}_i^*, i = 1, 2, \ldots, N_w\}$ to obtain the extracted watermark $\widetilde{W}$. 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 $\widetilde{W}$ is defined as

$$\text{NC}\left(W, \widetilde{W}\right) = \frac{\sum_{i=1}^{N_w} w_i \tilde{w}_i}{\sqrt{\sum_{i=1}^{N_w} w_i^2 \sum_i \tilde{w}_i^2}}, \tag{3}$$

where W and $\widetilde{W}$ denote an original watermark and extracted one, respectively, and $\widetilde{W} = \{\tilde{w}_i\}$ for $i = 1, 2, \ldots, N_w$. The watermark extracting process is shown in Figure 5.

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 4 optimal watermarking parameters. This allows the system to achieve optimal performance for digital audio watermarking.

For the optimization process, GA is applied in the watermark embedding and the watermark extracting processes to search for quantization steps (S_j). The objective function of searching process is computed by using factors that relate to both imperceptibility and robustness 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 Figure 6 and details of genetic algorithm optimization process are described as follows.

3.3.1. Chromosome Encoding. Chromosomes in GA represent desired parameter to be searched. The population (chromosomes) size must be chosen carefully to give sufficient genetic diversity. A small-sized population reduces the evaluation cost but it may cause premature convergence. For a large-sized population, the higher computation time is required. Therefore, it is important to achieve a good or acceptable result within a reasonable computation time. In this paper, the proper population size is set to be 40 (as described in the next section). The encoding scheme is binary string with 32 bit resolutions for each chromosome. Hence, the parameter S_j is represented by chromosome with length of 128 bits.
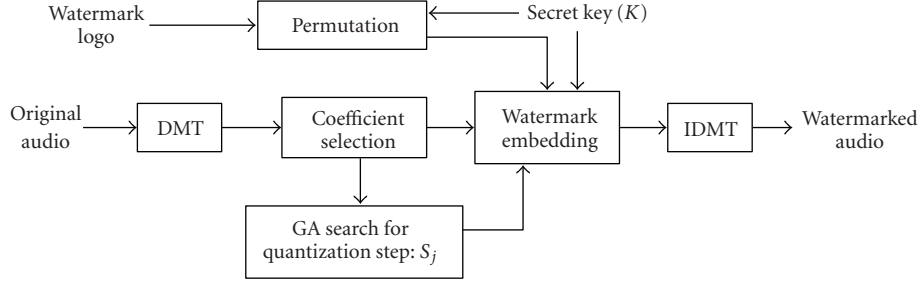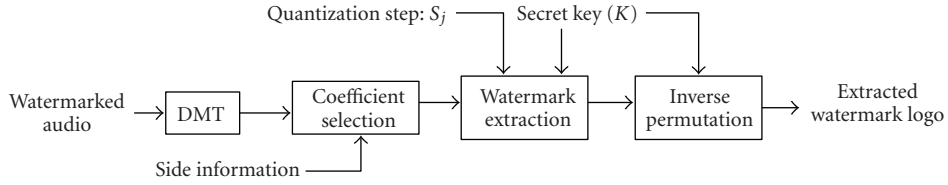
FIGURE 4: Watermark embedding process.

FIGURE 5: Watermark extracting process.

3.3.2. Objective Function Evaluation. The most critical step in the GA optimization process is the definition of a reliable objective function. Objective function is a fitness measure on solution represented by each chromosome. Its value tells how well the chromosome satisfies the final goal. In this paper, the objective function of GA is formed by combining two performance measures in digital audio watermarking, namely, average of the three normalized correlations (NC_{ave}) and difference (DIF) between desired SNR and obtained SNR from each iteration as performance indexes. Hence, DIF is an imperceptibility measure, while NC_{ave} is a robustness measure. The objective function (f_{obj}) can be computed as follow:

$$f_{obj} = \delta_{DIF} \times DIF + \delta_{NC} \times NC_{ave}, \qquad (4)$$

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.

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 robustness of the embedded watermark. They are MP3 compression at 64 kbps (Attack1), Gaussian noise addition (Attack2), and requantization (Attack3). 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 Figure 6.

TABLE 1: Parameters of genetic algorithm.

Parameters	Values
Population size	40
Chromosome length	128
Maximum number of generation	30
Crossover rate	0.7
Mutation rate	0.05

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_j can be optimally searched to achieve the best of both output audio quality and watermark robustness.

3.3.3. Selection, Genetic Operation, and Replacement. After evaluating fitness value of each chromosome based on the proposed objective function, chromosomes will be selected to produce offspring 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, that is, the quantization step S_j, is determined. Table 1 represents GA parameters used in the experiments. These parameters are referred to [22] and fixed by experiments.

TABLE 2: Average results for different population sizes.

| Population sizes | | Normalized correlation (NC) | | | f_{obj} |
		Attack1	Attack2	Attack3	
10	Average	0.999960	0.999970	0.999960	3.111
	Stdev	0.000084	0.000483	0.000069	
20	Average	0.999980	0.999970	0.999970	3.308
	Stdev	0.000632	0.000674	0.000048	
30	Average	0.999998	0.999960	0.999950	3.316
	Stdev	0.000031	0.000843	0.000084	
40	Average	0.999998	0.999998	0.999998	3.362
	Stdev	0.000031	0.000004	0.0000042	
50	Average	0.999960	0.999998	0.999998	3.368
	Stdev	0.000484	0.000031	0.000004	

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 7(a) and 7(b) show the original watermark and permuted watermark, respectively.

We use SNR, NC, and BER (Bit error rate) to analyze the performance of the proposed algorithm. The BER and SNR are defined as:

$$\text{BER} = \frac{\text{Number of error bits}}{\text{Number of total bits}} \times 100\%,$$

$$\text{SNR} = 10 \log_{10}\left(\frac{\sum_i (f_i^2)}{\sum_i (f_i - f_i')^2}\right), \qquad (5)$$

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

4.1. Results of Genetic Algorithm Optimization. The population size plays a crucial role in representing the solutions to the problem at hand. The results of varying the number of population are displayed in Table 2. The calculated average values (Average) and standard deviation (Stdev) values of the results of ten audio signals are also given in the tables. From the table, the objective value increase as the size of population. However, the differences between the objective values obtained from different population sizes are not very significant. Thus, the experiments were not carried out with the population size beyond 40.
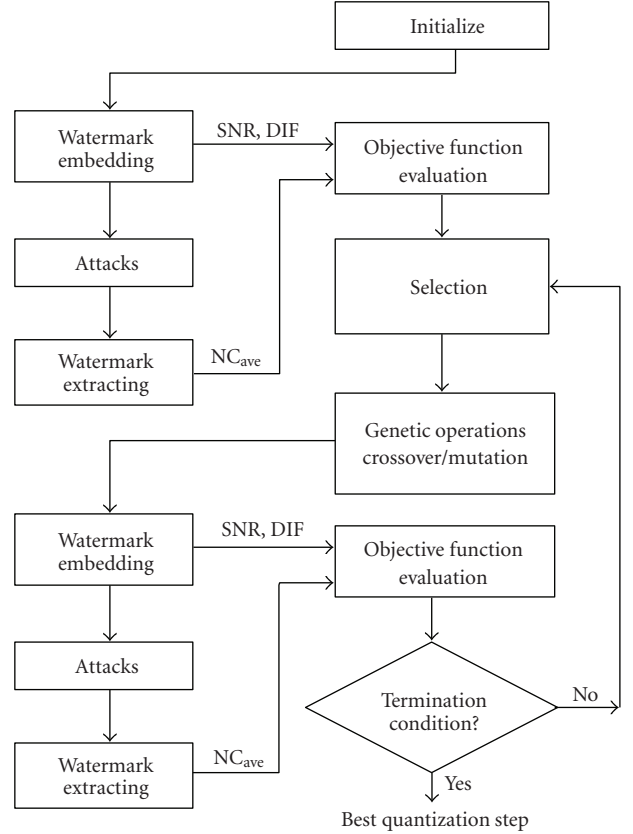

FIGURE 6: Optimization diagram for digital audio watermarking using genetic algorithm.
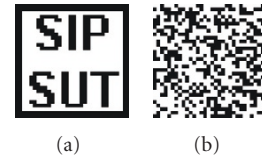

FIGURE 7: (a) Original watermark and (b) permuted watermark.

Figures 8 and 9 show the convergence of GA optimization at 30 generations of Classical1, and Rock1, respectively. It is obvious that as the number of generation increases, the objective function (f_{obj}) gradually approaches to a saturation value. The resulting parameters, which are 4 quantization steps from GA optimization of 10 test audios, are shown in Table 3. 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.

Figure 10 shows the original Classical1 audio signal waveform and the corresponding watermarked audio signal
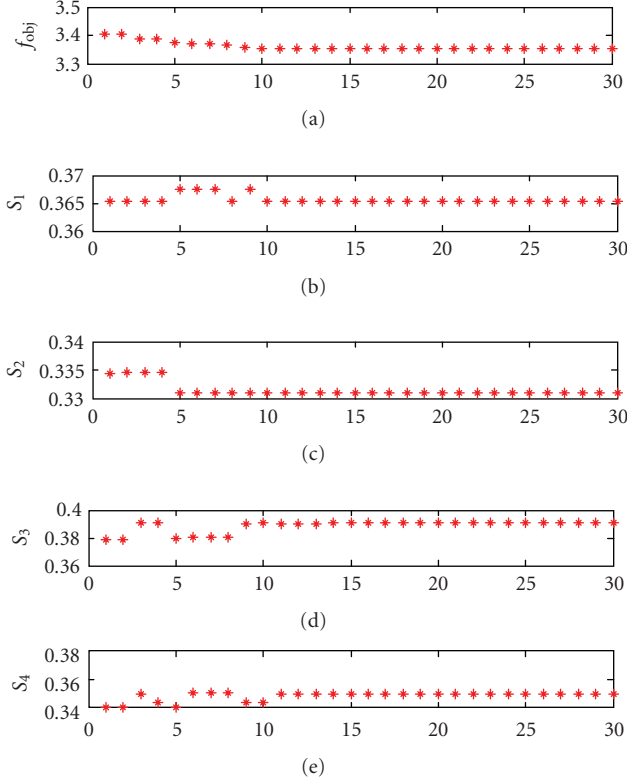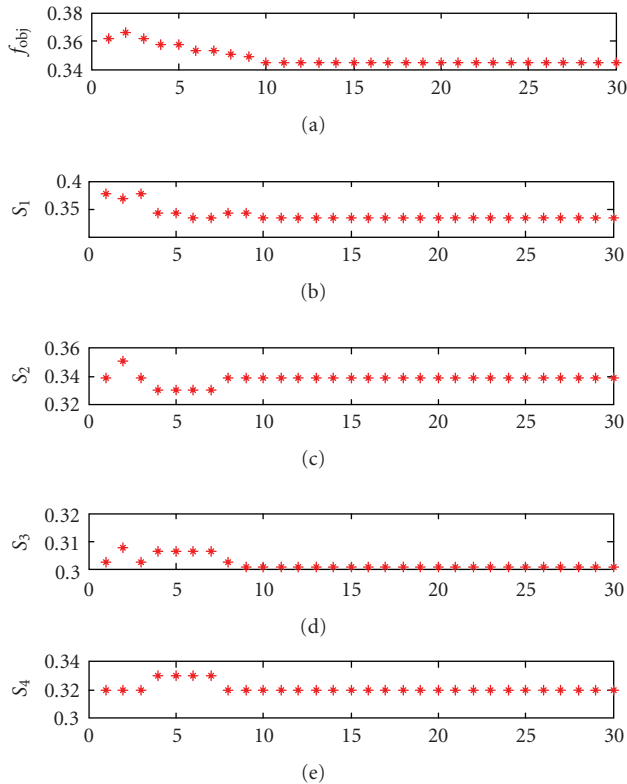
FIGURE 8: f_{obj}, S_1, S_2, S_3, and S_4 from GA optimization process (Classical1).

FIGURE 9: f_{obj}, S_1, S_2, S_3, and S_4 from GA optimization process (Rock1).

TABLE 3: Quantization steps from GA optimization of 10 test audios.

Host signal	Quantization step			
	S_1	S_2	S_3	S_4
Classical1	0.3654	0.3309	0.3905	0.3497
Classical2	0.3482	0.3391	0.3100	0.3196
Country1	0.3311	0.3402	0.3005	0.3215
Country2	0.3425	0.3308	0.3772	0.3211
Jazz1	0.3482	0.3398	0.3074	0.3024
Jazz2	0.3311	0.3960	0.3022	0.3084
Rock1	0.3340	0.3391	0.3008	0.3200
Rock2	0.3426	0.3342	0.3313	0.3115
Pop1	0.3311	0.3952	0.3005	0.3091
Pop2	0.3473	0.3477	0.3015	0.3194

TABLE 4: Signal-to-noise ratio of watermarked audio signals.

Host signals	SNR (dB)	
	Without-GA	With-GA
Classical1	25.31	26.04
Classical2	22.23	23.85
Country1	28.60	30.30
Country2	23.06	24.91
Jazz1	23.03	27.97
Jazz2	26.09	27.67
Rock1	26.73	28.30
Rock2	27.99	29.36
Pop1	21.82	24.90
Pop2	22.12	24.67
Average	24.69	26.79

waveform. Note that SNR is as high as 26.04 dB for watermarked audio signal. However, there is no obvious difference between original signal and watermarked signal by using informal listening test, and from Figures 10(a) and 10(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 4. 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 the 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.

4.3. Robustness Test Results. We first tested the robustness of the proposed algorithm to ten 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 eight types of attacks are then investigated.
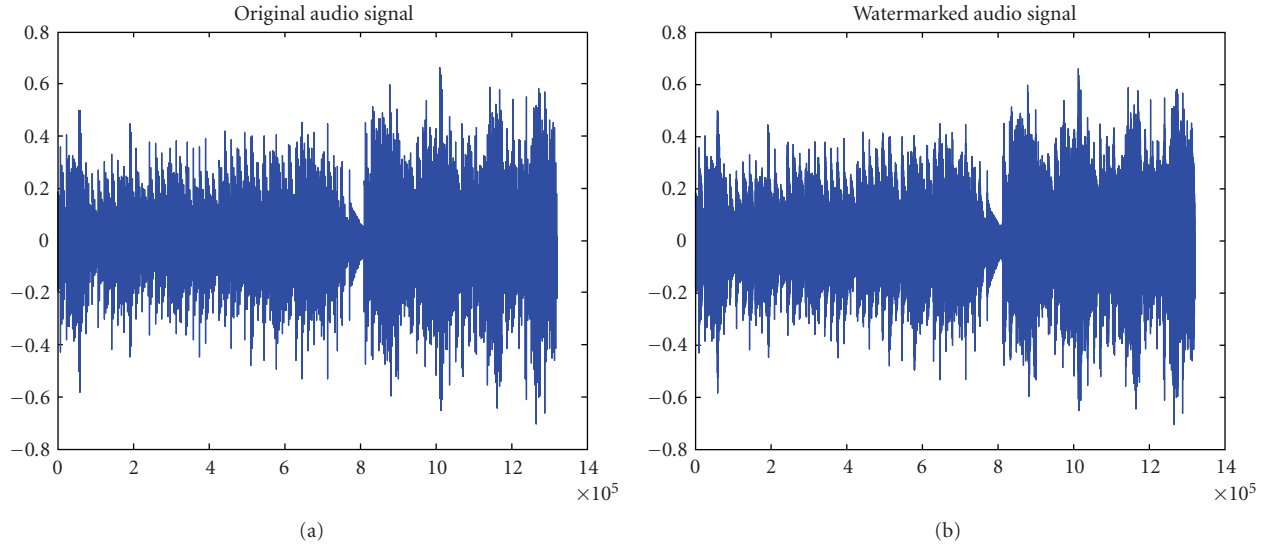
FIGURE 10: (a) Original audio signal (Classical1), (b) watermarked audio signal.

(1) Resampling: the audio signal is first downsampled at 22.05 kHz, and then upsampled at 44.1 kHz.

(2) Requantization: the 16-bit watermarked audio signals have been requantized 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 cutoff 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 bitrate 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.

(7) Digital-to-analog/analog-to-digital (DA/AD) with same rate reacquisition: the output of audio signal from a sound card is connected to the analog input of another computer.

(8) Time-scale modification (TSM): time-scale modification between +5% and −5% of the total audio excerpt length is performed.

Detection results for the various attacks described above are shown in Table 5 which displayed the NC and BER from watermark extraction. The experimental results given in Table 5 show that the watermark is not affected by resampling, requantization, 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, MP3 compression at 56 kbps, DA/AD, and TSM (±5%) attacks, the BER values of the recovered watermark sequence are 8.0133%, 5.2636%, 0.0977%, 17.2852%, 11.8164%, and 16.0616% for the Without-GA method and 6.9336%, 0.1953%, 0.0977%, 5.7619%, 2.2461%, and 2.1484% for the with-GA method, respectively. Although a lot of loss occurred in the audio

TABLE 5: Robustness comparison of our algorithm (average values).

Attack type	Without-GA		With-GA	
	NC	BER (%)	NC	BER (%)
Attack free	1.0000	0.0000	1.0000	0.0000
Re-sampling	1.0000	0.0000	1.0000	0.0000
Requantization	1.0000	0.0000	1.0000	0.0000
Low-pass filtering	0.9292	8.0133	0.9393	6.9336
Additive noise	1.0000	0.0000	1.0000	0.0000
Cropping	0.9031	5.2636	0.9983	0.1953
MP3-56 kpbs	0.9992	0.0977	0.9992	0.0977
MP3-64 kpbs	1.0000	0.0000	1.0000	0.0000
MP3-96 kpbs	1.0000	0.0000	1.0000	0.0000
MP3-128 kpbs	1.0000	0.0000	1.0000	0.0000
DA/AD	0.8486	17.2852	0.9488	5.7617
TSM +5%	0.8970	11.8164	0.9805	2.2461
TSM −5%	0.8660	16.0616	0.9814	2.1484

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.

Then, a rough comparison is given in Table 6 based on embedding data payload and BER under MP3 compression with bit rates of 64 kbps and 128 kbps. It can be seen that the data payload of the proposed algorithm is significantly higher than those of [5, 7, 10]. As shown in Table 6, the detected BER values are all zero which indicates that compression and decompression have no effect on 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 [10]. In order to compare robustness between the two techniques

Table 6: Algorithm comparison.

Algorithms	Data payload (bps)	BER under MP3 compression (64 kbps)	BER under MP3 compression (128 kbps)
With-GA	34.14	Approximately 0.00%	Approximately 0.00%
[5]	8.54	Approximately 2.99%	Not available
[7]	10.72	Not available	Approximately 3.56%
[10]	21.00	Approximately 5.3722%	Approximately 3.6133

Table 7: Algorithm comparison.

	Attack free	Re-sampling	Requantization	Low-pass filtering	Additive noise
[10]	SIP SUT	SIP SUT	SIP SUT	SIP SUT	SIP SUT
BER (%)	0	0	0.1953	3.3203	2.6367
SNR (dB)	24.1209	23.5533	20.0936	10.3506	19.5607
With-GA	SIP SUT	SIP SUT	SIP SUT	SIP SUT	SIP SUT
BER (%)	0	0	0	1.1719	0
SNR (dB)	24.5518	23.5052	20.9320	10.3884	19.5684
	Cropping	MP3 128 kbps	MP3 96 kbps	MP3 64 kbps	MP3 56 kbps
[10]	SIP SUT	SIP SUT	SIP SUT	SIP SUT	SIP SUT
BER (%)	5.3711	3.6133	4.1016	5.3711	5.7617
SNR (dB)	10.8289	19.5607	19.5537	18.5348	18.5400
With-GA BER (%) SNR (dB)	SIP SUT	SIP SUT	SIP SUT	SIP SUT	SIP SUT
	0.1953	0	0	0	0
	16.3075	19.7552	19.5920	18.8265	18.6666

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 7.

Table 7 shows test results of Rock1 audio signal with no attack, re-sampling, requantization, low-pass filtering, addition of noise, cropping, and MPEG 1 Layer III compression with compression rates of 128, 96, 64 kbps and 56 kbps, respectively. The BER of watermark signal and the SNR of digital audio signal are also displayed.

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 [10]. The results demonstrate that our proposed method yields significantly more robust watermark than the method in [10] does.

5. Conclusions

This paper proposes a digital 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 multiresolution 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. Performance improvement with respect to existing algorithms is obtained by genetic algorithm optimization. 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 and psychoacoustic model.

Acknowledgments

References

[1] S.-J. Lee and S.-H. Jung, "A survey of watermarking techniques applied to multimedia," in *Proceedings of the IEEE International Symposium on Industrial Electronics*, vol. 1, pp. 272–277, Pusan, South Korea, June 2001.

[2] S. Katzenbeisser and F. A. P. Petitcolas, *Information Hiding Techniques for Steganography and Digital Watermarking*, Artech House, Norwood, Mass, USA, 2000.

[3] I. J. Cox, J. Kilian, F. T. Leighton, and T. Shamoon, "Secure spread spectrum watermarking for multimedia," *IEEE Transactions on Image Processing*, vol. 6, no. 12, pp. 1673–1687, 1997.

[4] C.-H. Huang and J.-L. Wu, "Watermark optimization technique based on genetic algorithms," in *Security and Watermarking of Multimedia Contents II*, vol. 3971 of *Proceedings of SPIE*, pp. 516–523, San Jose, Calif, USA, January 2000.

[5] J. W. Seok and J. W. Hong, "Audio watermarking for copyright protection of digital audio data," *IEE Electronics Letters*, vol. 37, no. 1, pp. 60–61, 2001.

[6] I. Orović, P. Zogović, N. Žarić, and S. Stanković, "Speech signals protection via logo watermarking based on the time-frequency analysis," *Annals of Telecommunications*, vol. 63, no. 7-8, pp. 369–377, 2008.

[7] R. Wang, D. Xu, J. Chen, and C. Du, "Digital audio watermarking algorithm based on linear predictive coding in wavelet domain," in *Proceedings of the 7th International Conference on Signal Processing (ICSP '04)*, vol. 3, pp. 2393–2396, Beijing, China, August 2004.

[8] J. Huang, Y. Wang, and S. Wu, "Audio watermarking scheme robust against desynchronization based on the dyadic wavelet transform," *EURASIP Journal on Advances in Signal Processing*, vol. 2010, Article ID 232616, 17 pages, 2010.

[9] S. Wu, J. Huang, D. Huang, and Y. Q. Shi, "Efficiently self-synchronized audio watermarking for assured audio data transmission," *IEEE Transactions on Broadcasting*, vol. 51, no. 1, pp. 69–76, 2005.

[10] N. Sriyingyong and K. Attakitmongcol, "Wavelet-based audio watermarking using adaptive tabu search," in *Proceedings of the 1st International Symposium on Wireless Pervasive Computing*, vol. 1, pp. 1–5, Phuket, Thailand, January 2006.

[11] X.-Y. Wang, P.-P. Niu, and H.-Y. Yang, "A robust, digital-audio watermarking method," *IEEE Multimedia*, vol. 16, no. 3, pp. 60–69, 2009.

[12] R.-D. Wang, D.-W. Xu, and Q. Li, "Audio watermarking algorithm based on wavelet packet and psychoacoustic model," in *Proceedings of the 6th International Conference on Parallel and Distributed Computing, Applications and Technologies (PDCAT '05)*, vol. 1, pp. 812–814, Dalian, China, December 2005.

[13] K. R. Kwon and A. H. Tewfik, "Adaptive watermarking using successive subband quantization and perceptual model based on multiwavelet transform," in *Security and Watermarking of Multimedia Contents IV*, vol. 4675 of *Proceedings of SPIE*, pp. 334–348, San Jose, Calif, USA, January 2002.

[14] P. Kumsawat, K. Attakitmongcol, and A. Srikaew, "A new approach for optimization in image watermarking by using genetic algorithms," *IEEE Transactions on Signal Processing*, vol. 53, no. 12, pp. 4707–4719, 2005.

[15] L. Ghouti and A. Bouridane, "A robust perceptual audio hashing using balanced multiwavelets," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP '06)*, vol. 1, pp. 209–212, Toulouse, France, May 2006.

[16] 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.

[17] F. Hartung and M. Kutter, "Multimedia watermarking techniques," *Proceedings of the IEEE*, vol. 87, no. 7, pp. 1079–1107, 1999.

[18] S. Sedghi, H. R. Mashhadi, and M. Khademi, "Detecting hidden information from a spread spectrum watermarked signal by genetic algorithm," in *Proceedings of the IEEE Congress on Evolutionary Computation (CEC '06)*, vol. 1, pp. 173–178, Vancouver, Canada, July 2006.

[19] K. Attakitmongcol, D. P. Hardin, and D. M. Wilkes, "Multiwavelet prefilters—Part II: optimal orthogonal prefilters," *IEEE Transactions on Image Processing*, vol. 10, no. 10, pp. 1476–1487, 2001.

[20] J. S. Geronimo, D. P. Hardin, and P. R. Massopust, "Fractal functions and wavelet expansions based on several scaling functions," *Journal of Approximation Theory*, vol. 78, no. 3, pp. 373–401, 1994.

[21] J. H. Holland, *Adaptation in Natural and Artificial Systems*, The University of Michigan Press, Ann Arbor, Mich, USA, 1975.

[22] D. E. Goldberg, *Genetic Algorithms in Search Optimization and Machine Learning*, Addison-Wesley, Reading, Mass, USA, 1989.