Intelligent Audio Watermarking using Genetic Algorithm in DWT Domain

M. Ketcham, and S. Vongpradhip

Abstract**—**In this paper, an innovative watermarking scheme for audio signal based on genetic algorithms (GA) in the discrete wavelet transforms is proposed. It is robust against watermarking attacks, which are commonly employed in literature. In addition, the watermarked image quality is also considered. We employ GA for the optimal localization and intensity of watermark. The watermark detection process can be performed without using the original audio signal. The experimental results demonstrate that watermark is inaudible and robust to many digital signal processing, such as cropping, low pass filter, additive noise.

Keywords**—**Intelligent Audio Watermarking, Genetic Algorithm, DWT Domain.

I. Introduction

CURRENTLY, multimedia data have become wide-spread via the internet. It may be that the ease with which perfect copies can be made will lead to large-scale unauthorized. There has therefore been significant recent research into 'steganography' to copyright, such as Digital watermarking which has consequently been introduced as a complementary Copyright protection technology. The enormous growth of the digital world: Old analog audio tapes were substituted by digital disks, personal computers with internet connections took homes by storms and digital versatile disk (DVD) players invaded living rooms. Unfortunately, this has also raised many concerns regarding copyright protection since digital data can be perfectly duplicated and rapidly redistributed on a large scale. Today, even nontechnical users can exchange copyrighted material via peer-to-peer networks and multimedia content providers have requested security mechanisms before releasing their highly valued property. Many digital right management (DRM) frameworks rely on end-to-end encryption to make digital data completely unusable without the proper decryption key. However, this protection falls when encrypted data is decrypted to eventually be presented to a human user. Digital watermarking was consequently introduced in the 1990s as a second line of defense to fill this *analog hole*. There are many techniques of audio watermarking algorithms can be grouped into third categories: patchwork in the frequency domain [1], echo hiding in the time domain [2, 3, 4, 5, 6 and 7], and spread-spectrum [8, 9, 10, 11, 12, 13, 14, 15, 16 and 25]. To improve the performance of watermarking schemes is to make use of artificial intelligence techniques. The image watermarking problem can be viewed as an optimization problem. Therefore, it can be solved by genetic algorithms (GA). There has been little research in application of GA to image watermarking problems. Huang et al. [17, 18] proposed a watermarking method based on the DCT and GA. They embedded the watermark with visually recognizable patterns into the image by selectively modifying the middle-frequency parts of the image. The GA is applied to search for the locations to embed the watermarking in the DCT coefficient block such that the quality of the watermarked image is optimized. Shieh et al. [19] presented a watermarking optimization technique similar to [17, 18]. They use GA to find the optimum frequency bands for watermark embedding into a DCT-based watermarking system, which can simultaneously improve security, robustness, and the image quality of the watermarked image. Genetic algorithm (GA) has established itself as one of the most powerful and applicable optimization optimization methods. Because of its high flexibility and robustness beside its simplistic implementation procedure, the GA has been employed for a large number of applications in signal processing areas as a powerful optimization tool. Successful operation of GA is widening its application in signal processing.

In this paper, we present a new method for audio watermark systems. Our paper is based on the methodology: in order to manipulate search for optimal localization. This is the idea of utilizing the evolutionary algorithms in the field of audio watermark systems.

The rest of the paper is organized as follows. In section 2, we briefly review the watermarking schemes, discrete wavelet transforms (DWT), data payload and genetic algorithm. In section 3, we explain the methodology to incorporate genetic algorithms for the embedding algorithm. The detection algorithm is shown in section 4, and we conclude this paper in section 5.

Manuscript received March 29, 2007. This work was supported in part by the Department of Computer Engineering Faculty of Engineering , Chulalongkorn University.

M. Ketcham is with Department of Computer Engineering Faculty of Engineering, Chulalongkorn University (e-mail: maoquee@hotmail.com).

S. Vongpradhip, is with Department of Computer Engineering Faculty of Engineering, Chulalongkorn University (e-mail: sartid@cp.eng.chula.ac.th).

II. Notation and Background

A. Watermarking Schemes

The basic idea of watermarking is to add a watermark signal into the host data to be watermarked such that the watermark signal is unobtrusive and secure in the signal mixture, but can partly or fully be recovered from the signal mixture later on if the correct cryptographically secure key is used. In order to ensure imperceptibility of the modification caused by watermark embedding, we use a perceptibility criterion of some sort. This can be implicit or explicit, fixed or adaptive to host data. **As** a consequence of the required imperceptibility, the individual samples (e.g., pixels or transform coefficients) that are involved in watermark embedding can only be modified by an amount relatively small to their average amplitude. In order to ensure robustness despite of the small allowed changes, we usually distribute the watermark information redundantly over many samples (e.g., pixels) of the host data. Therefore, the recovery is more robust if more watermarked data are available for recovery.

1) Patchwork Scheme

The patchwork scheme [1] embeds a special statistic into an original audio signal. The two major steps in the scheme are: (i) choose two patches pseudo-randomly and (ii) add the small constant value d to the samples of one patch A and subtract the same value d from the samples of patch B. Mathematically speaking

$$a_i^* = a_i + d; b_i^* = b_i - d \tag{1}$$

Where a_i and b_i are samples of the patchwork sets a and b. Thus, the original sample values have been slightly modified.

2) Echo Hiding Scheme

Echo hiding [2, 3, 4, 5, 6, 7] is a method for embedding information into an audio signal. It seeks to do so in a robust fashion, while not perceivably degrading the original signal. Echo hiding has applications in providing proof of the ownership, annotation, and assurance of content integrity. Therefore, the embedded data should not be sensitive to removal by common transform to the embedded audio, such as filtering, re-sampling, block editing, or lossy data compression. Echo hiding embeds data by introducing an echo. The value of a hidden datum corresponds to the time delay of the echo (for example, d for 1 and d'for 0) and its amplitude. The echo delays are selected to be less than the detectable hearing limit.

$$y[n] = x[n] + \alpha[n - \delta] \tag{2}$$

Where $x[n]$ is the original audio signal, α is the echo's amplitude and δ is the audio signal delay, the delay time need to be below the threshold at the human ear can resolve the echo. The data are hidden by varying two parameters of the echo: initial amplitude, and offset. As the delay between the original and echo decreases, the two signals blend.

3) Spread-Spectrum

Spread-spectrum watermarking scheme is an example of the correlation method which embeds pseudo-random sequence and detects watermark by calculating correlation between pseudo-random noise sequence and watermarked audio signal. This scheme spreads pseudo-random sequence across the audio signal (Fig. 1).

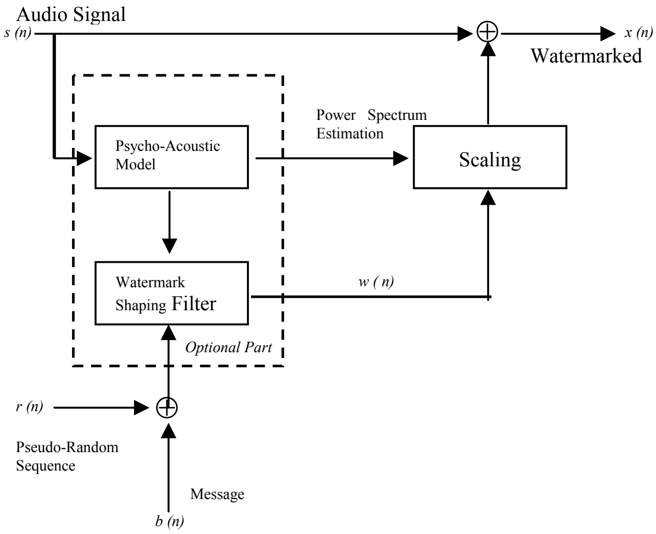

Fig. 1 A typical embedded of the spread-spectrum watermarking scheme

The wideband noise can be spread into either time domain signal or transform-domain signal no matter what transform is used. Frequently used transforms include DCT (Discrete Cosine Transform), DFT (Discrete Fourier Transform), and DWT (Discrete Wavelet Transform). The binary watermark message $\upsilon = \{0,1\}$ or its equivalent bipolar variable *b*= {-1, +1} is modulated by a pseudo-random sequence *r (n)* generated by means of a secret key. Then the modulated watermark *w (n) = br (n) is* scaled according to the required energy of the audio signal *s (n)*. The scaling factor α controls the trade-off between robustness and inaudibility of the watermark. The modulated watermark *w (n)* is equal to either *r (n)* or *-r (n)* depending on whether *v = 1* or *v = 0.* The modulated signal is then added to the original audio to produce the watermarked audio *x (n)* such as

$$x(n) = s(n) + \alpha w(n) \tag{3}$$

B. Discrete Wavelet Transform (DWT)

The discrete wavelet transform has received a tremendous amount of interest in many important signal processing applications including audio and image watermarking [20, 21, and 22]. It has been developed with the idea of looking at a signal at various scales and analyzing it with various resolutions. The basis functions are obtained from a single prototype wavelet by dilations, contractions and shifts. The principle objective of the wavelet transform is to hierarchically decompose an input signal into a series of successively lower frequency approximation sub band and their associated detail sub bands. For the dyadic wavelet decomposition, at each level, the low frequency approximation sub band and detail sub band (or sub bands for multidimensional case) contain the information needed to reconstruct the low frequency approximation signal at the next higher resolution level. Wavelet techniques provide excellent space and frequency energy compaction, in which energy tends to a cluster spatially in each subband. For DWT, the link between the spatial/temporal domain signals, f (t), and the DWT of f (t), d (k,l),is

$$f(t)=\sum_{k=-\infty}^{\infty}\sum_{l=-\infty}^{\infty}d(k,l)2^{-k/2}\psi(2^{-k}t-l). \qquad (4)$$

Where ψ (•) denotes the mother wavelet.

C. Data Payload

The *data payload* refers to the number of bits that are embedded into origin audio within a unit of time, measured with *bps* (bit persecond). Suppose that the sampling rate of audio is R (Hz) persecond and the wavelet decomposition level is K. The data payload B of the algorithm is defined as:

$$B = R/2^k\ bps \qquad (5)$$

D. Genetic Algorithm

GA is a search technique for finding the global maximum/minimum solutions for problems. Although the GA operation performs randomly, choosing candidates to avoid stranding on a local optimum solution, there is no guarantee that the global maximum/minimum will be found. In general, the possibility of obtaining the global maximum/minimum by using GA is related to the complexity of a problem. That is the more complex a problem is, the higher the difficulty of obtaining the optimum solution [23]. The GA is one of the most widely used artificial intelligent techniques for optimization. They have been successfully applied to obtain good solutions in optimal localization and intensity of audio watermark. Usually, the GA starts with some randomly selected genes as the first generation, called *population.* Each individual in the population corresponding to a solution in the problem domain is called *chromosome.* An objective, called *fitness function,* is used to evaluate the quality of each chromosome. The chromosomes of high quality will survive and form a new population of the next generation. By using the tree operators: selection crossover, and mutation, we recombine a new generation to find the best solution. In order to apply the GA for embedding audio watermarking into the DWT the chromosomes is used to adjust position values of audio watermarking on DWT.

III. Embedding Algorithm

In this paper, the watermark data is a binary logo image. The embedding algorithm is performed to the wavelet coefficients obtained from 5-level wavelet decomposition. The security of the algorithm is enhanced by performing a random permutation to the watermark image. The watermarking embedding algorithm is as follows:

1. The watermark data, which is a M_1 x M_2 binary logo image, is transformed into a unidimentional antipodal sequence $w(i) \in \{+1, -1\}$ where M_1 and M_2 are the number of rows and columns of the binary watermark image. Then, we generate the random sequence $r(i)$ which is used to encrypt watermark to ensure security.

2. The input audio signal sampled at 44100 Hz is decomposed into five levels using 4-coefficient Daubecies wavelet (Db4). Next, all obtained wavelet coefficients at coarsest approximation subband are divided into k segments where $k=M_1M_2$. After this step, we have segmented wavelet coefficients, $C_k(i)$.

3. The average value of each segment $C_k(i)$, $m_k(i)$, is calculated and removed from all wavelet coefficients in $C_k(i)$ at the coarsest approximation subband to facilitate the embedding process. Let $C_k'(i)$ be the modified $C_k(i)$ by

$$C_k^{'}(i)=C_k(i)-m_k(i) \qquad (6)$$

4. Then, each bit of watermark data is embedded into each $C_k'(i)$ using the following method: if $w(i)=1$, all coefficients in selected segment are added by If $w(i) = -1$, they are subtracted β is the magnitude of $m_k(i)$

$$y_k(i)=C_k^{'}(i)+\beta w(i) \qquad (7)$$

5. IDWT is applied to the modified wavelet coefficients to transform them back to the audio signal in time domain.

The steps for applying GA into the watermarking system can be summarized embedding algorithm is shown in Fig. 2.

A. Various Attacks

Filtering To test the robustness against filtering a bandpass filter was applied to the watermarked signal by amplifying the signal by -9 dB in the low and high frequency domain. The cut-off frequency has been 441 Hz for the low-pass and 4410 Hz for the high-pass filter. These filtering results in audible distortions in comparison to the original signal.the paper will be sent. Proofs are sent to the corresponding author only.

Robustness against cropping The embedding of the watermark in every time slice of about 1.2 seconds enables the detection of the watermark even in the case of cropping or cutting. The necessary precondition for reading the watermark is to have a contiguous part of at least about 2.5 seconds from the watermarked audio stream.

Random noise: Addition of random numbers to the samples, constrained by a parameter giving the relative amount of the number compared with the original signal. Up the 0,91 % of the original sample value could be added as noise without degrading the perceived quality.

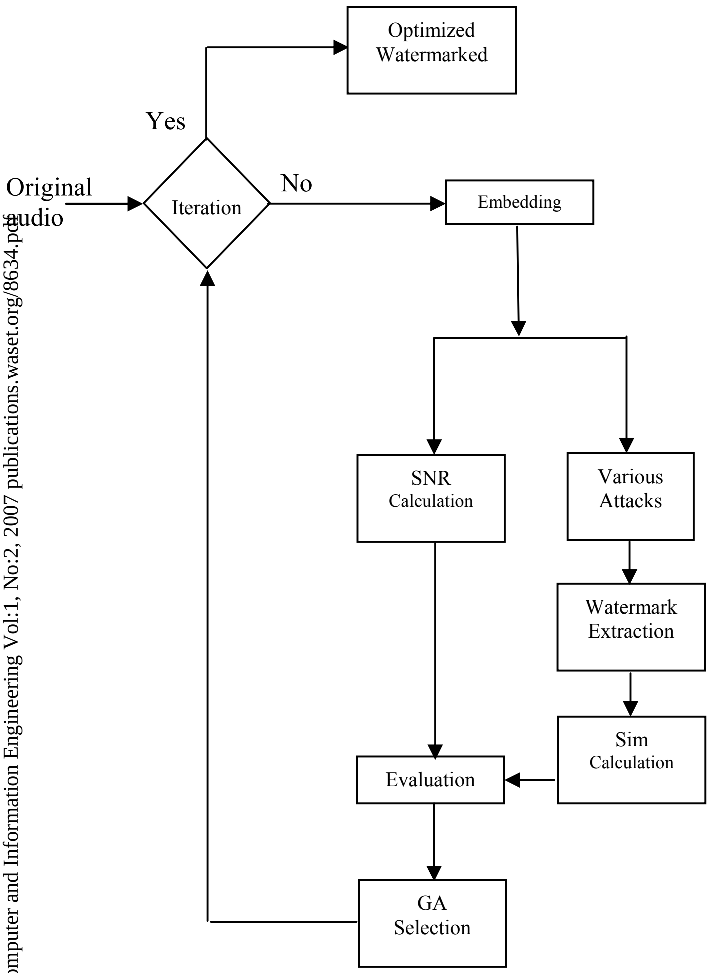

Fig. 2 The block diagram for watermark embedding with GA

B. Performance

To demonstrate the performance of this scheme, we used test watermark image is a binary image with the size of 25 x 25 pixels, binary-valued, see in Fig. 3. as the watermark.

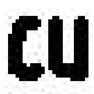

Fig. 3 The watermark used in experiments.

We applied the signal-to-noise-ratio (SNR) for evaluating the quality of the watermarked image, and similarity (Sim) to evaluate the robustness. The definitions of SNR and Sim are:

$$SNR = 10\log_{10}\left(\frac{\sum_i (Y_i^2)}{\sum (Y_i - y_i)^2}\right) \tag{8}$$

Where Y and y are audio signal original and audio signal embedding respectively.

$$Sim = \frac{\sum_{i=1}^{M1}\sum_{j=1}^{M2} w(i,j)w^*(i,j)}{\sqrt{\sum_{i=1}^{M1}\sum_{j=1}^{M2} W(i,j)^2}\sqrt{\sum_{i=1}^{M1}\sum_{j=1}^{M2} W^*(i,j)^2}} \tag{9}$$

Where W and W^* are original and extracted watermarks, respectively, i and j are indexes of the binary watermark image.

C. Watermarking Scheme with GA

In this section, we will introduce how to employ genetic algorithm into the mentioned DWT watermarking scheme. We will define an object function and use GA to optimize this function. The diagram of our proposed algorithm is shown in Fig. 2, and details of the GA are described as follows.

```
Pmut = 0.3;
 Pcross = 0.9;
 Popsize = 500;
Generation = 100;
Initialization (pop1 ,..., pop1000);
 Evaluation (pop1 ,..., pop1000)  /* Embedding ();
                                  Detection ();
              Various attacks ();
                        */
For loop = 1 to generation do
         Selection (tournament selection size 2);
         Mutation ();
        Evaluation (pop1 ,..., pop1000)
End For
```

Fig. 4. A pseudo-code for GA Algorithm

1) Chromosome Encoding

There are 100 chromosomes used in this work. The chromosomes are encoded using a 1000-digit binary string encoding scheme. Each chromosome is represented by the positions of embedding. There are 625 ones (1's) and 375 zeros (0's), where the ones are the position of the watermarks.

2) Objective Function Evaluation

The objective function of GA is composed of the signal-to-noise-ratio (SNR) of the watermarked signal versus the host signal used as a quality measure and the similarity (Sim) between extracted watermark and original watermark. The objective value f_i can be computed by the following equation:

$$f_i = SNR_i + \frac{1}{p}\sum_{n=1}^{p}(Sim_{n,i}) \qquad (10)$$

where p is the number of attacking schemes.

3) Operator

Selection: The selection method chosen is tournament selection with both tournament size and tournament probability as parameters to be optimized. The best chromosome in a tournament is determined by fitness, as described above.

Crossover: Crossover needs to be redesigned. The redesigned crossover has to keep the number of 1 bits (in each chromosome) to be stable at 625 and equal to those of the watermark, and the crossover rate is equal to 0.9. The process of crossover is as follow (see Figure 4):

Step 1: Randomly generate two chromosomes.
Step 2: Search positions that two chromosomes are different (Set P).
Step 3: Randomly select two positions from the set P
Step 4: Consider each chromosome. If the two positions selected in Step 3 are different, swap the genes in those positions.

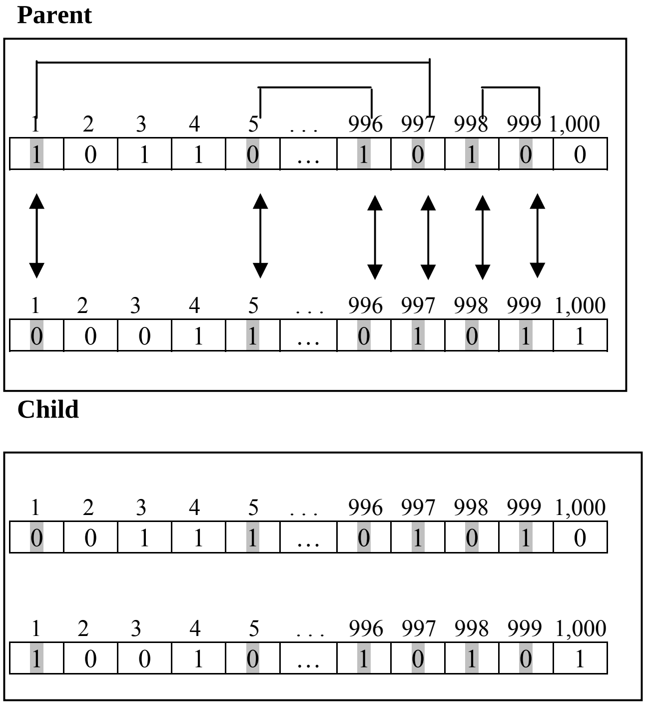

Fig. 5 An example for crossover operation

Mutation: The reverse operator method works by dividing a bit string into sections of size N. After that, the needed sections are chosen and mutated within that part.

IV. Detection Algorithm

The detection algorithm is performed without using the original audio signal. We first decompose watermarked audio signal into 5-level wavelet decomposition. Then, we segment the coefficients at the coarsest approximation sub-band as in the embedding process and calculate the mean value of each segment of audio signal wavelet coefficients. If the mean is larger than zero, a bit "1" is detected. If the mean is lower than zero, a bit "-1" is detected. This step is repeated until all embedded bits are detected. Then, we decrypt the watermark by using the same random sequence used in embedding procedure. Finally, all detected bits are rearranged to form a binary image as a detected watermark.

V. Experimental Result

In order to verify the performance, we test our algorithm on 8-bit signed mono audio signals sampled at 16 kHz with the length of about 18 seconds in WAVE format. According to the International Federation of the Phonographic Industry (IFPI), the SNR of the watermarked audio signal should be greater than 20 dB [22, 24].

TABLE I
The SNR and Sim Values Under Different GA Iterations in Eg .10

Number of generation	SNR (DB)	Sim_1	Sim_2	Sim_3
0	19.251	0.912	0.920	0.940
25	21.32	0.962	0.935	0.954
50	23.358	0.970	0.957	0.967
75	23.932	0.940	0.95	0.95
100	25.414	0.97	0.96	0.97

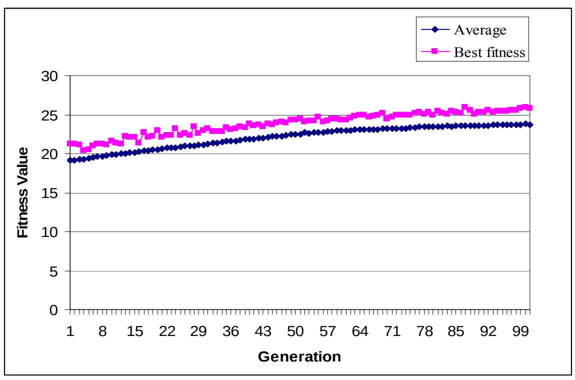

Fig. 6 Fitness evolution

From the data in Table I, we find that both the SNR and Sim values increase in generation. We also represent the relationships among the watermarked image quality and extract the Sim values with the number of generation. A population of 500 individuals (m=0.3) is used. In this example, the genetic algorithm was run for 100 generations. Fig. 6 shows the evolution of chromosome length and average best fitness.

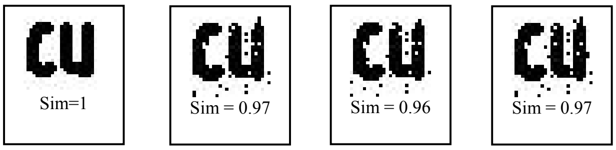

Fig. 7 Extracted watermark image with various similarity values

Extracted watermark images with difference similarities are shown in Fig. 7. The comparison of our algorithm is made with the algorithm using GA

VI. Conclusion

In this paper, we proposed an audio watermarking and robust algorithm using DWT-based GA. It is robust because we make use of GA to the evaluation selection of the embedding regions. In addition to the robustness of the proposed algorithm, we also improve the watermarked image quality with the aid of GA. Furthermore, the proposed algorithm extends to the application for the progressive audio watermarking scheme.

Experimental result reveal that if we just borrow the concepts of existing algorithm, both the watermarked image quality and the Sim value use of the extracted watermarks after certain attacks will be poor. We also give the result of the progressive audio watermarking scheme to show the effectiveness of watermarking based on GA.. The GA offers a systematic way to consider the improvements of the evaluation. With the experimental result under a variety of attacking, we are able to require its robustness and greatness over the existing algorithm with the scheduled techniques. Here we show that the performance of the audio watermarking can be further enhanced by the combination of quality and robust.

To the knowledge of the authors, this is the application of GA for audio watermarking and widens application of GA in the area of signal processing. Possible future works of the proposed approach will aim at making suitable enhancements for genetic watermarking. Firstly, the GA algorithm should be able to search for the optimal intensity of watermarks. Secondly, the algorithm should support parallel GA audio watermarking as well.

References

[1] Yeo, I. K., and Kim HJ, "Modified patchwork algorithm: A novel audio watermarking scheme", *IEEE Transactions on Speech and Audio Processing*, vo1-11, 2003.
[2] W. Bender, D.Gruhl, and N. Morimoto, "Techniques for data hiding" *IBM Systems journal,* 1996, 35(3/4): 131-336.
[3] M.Cooperman and S.Moskowitz, "Steganographic method and device", *U.S. Patent* 5, 913,004, Mar., 1997.
[4] J. Huang, Y. Wang, and Y.Q. Shi, "A blind audio watermarking algorithm with self-synchronization", *Proc. Of IEEE, Int. Sym. On Circuits and Systems*, vol. 3 pp. 627-630, 2002.
[5] D. Gruhl, A. Lu and W. Bender, "Echo Hiding", *Proc. Of 1st Information Hiding Workshop*, LNCS vol. 1174, Berlin, Germany: Springer – Verlag, pp. 295 -315, 1996.
[6] P. Bassia, I. Pitas, and N. Nikoladis, "Robust audio watermarking in time domain", *IEEE Transactions on Multimedia*, vo1. 3, pp 232-241, 2001.
[7] Kim, H.J., and Choi Y.H., "A novel echo-hiding scheme with backward and forward kernels", *IEEE Transactions on circuits and Systems for video and technology*, vo1. 13, 2003.
[8] L. Cui, S. Wang and T. Sun, "The application of binary image in digital audio watermarking," *Neural Networks and Signal Processing,* vol 2, pp. 1497 – 1500, 2003.
[9] Sang-Kwang, Lee and Yo Sung-Ho, "Digital audio watermarking in the cepstrum domain", *IEEE Transaction* on Consumer Electionics, vo1. 46 No. 3, 2000.
[10] Boney, L., Tewfik, A. H., and Hamdy, K. N., "Digital watermarks for audio signal," *International Conference on Multimedia Computing and Systems*, pp. 473-480, 1996.
[11] Cox, I.J., Kilian, J., Leigton, F.T., and Shamoon, T., "Secure spread spectrum watermarking for multimedia," *IEEE Transactions on Image Processing*, vol. 6, pp. 1673-1687, 1996.
[12] Cvejic, N., Keskinarkaus, A., and Seppänen, T., "Audio watermarking using m-sequences and temporal masking," *IEEE Workshops on Applications of Signal Processing to Audio and Acoustics*, pp. 227-230, 2001.
[13] Kim, H., "Stochastic model based audio watermark and whitening filter for improved detection," *IEEE International to Conference on Acoustics, Speech, and Signal Processing*, vol. 4, pp. 1971-1974, 2000.
[14] Kirovski, D. and malvar, H., "Robust spread-spectrum audio watermarking," *IEEE International Conference on Acoustics, Speech, and Signal Processing*, pp. 1345-1348, 2001.
[15] Seok, J., Hong, J., and Kim, J., "A novel audio watermarking algorithm for copyright protection of digital audio," *ETRI Journal*, vol. 24, pp. 181-189, 2002.
[16] Swanson, M., Zhu, B., Tewfik, A., and Boney, L., "Robust audio watermarking using perceptual masking," *Signal Processing*, vol. 66, pp. 337-355, 1998.
[17] C.-H. Huang and J.-L.Wu, "A watermark optimization technique based on genetic algorithms," in *Proc. SPIE—Visual Communications ImageProcessing*, vol. 3971, Feb. 2000, pp. 516–523.
[18] -----, (2004) Using genetic algorithms as watermarking performance optimizers. [Online]. Available: http://www.cmlab.csie.ntu.edu.tw/~bh/oldpage/
[19] C.-S. Shieh, H.-C. Huang, F.-H. Wang, and J.-S. Pan, "Genetic watermarking based on transform domain techniques," *Pattern Recognit.*, vol. 37, no. 3, pp. 555–565, Mar. 2004.
[20] Michael A., Gerzon and peter G. Graven., "A high-rate buried-data channel for audio CD", *Journal of the Audio Engineering Society,* 43(1/2):3-22, January – February 1995.
[21] Shaoquan Wu, J. Huang, Daren Huang, Y.Q. Shi, "Self-Synchronized Audio Watermark in DWT Domain", *IEEE (ISCAS)*, v-712 –v-715, 2004.
[22] N. Sriyingyong, and K. Attakitmongcol "Wavelet-Based Audio Watermarking Using Adaptive Tabu Search" *Wireless Pervasive Computing, 2006 1st International Symposium on* 16-18 an. 2006 Page(s):1 – 5, 2006.
[23] J. H. Holland, *Adaptation in Natural and Artificial Systems.* Ann Arbor: Univ. of Michigan Press, 1975.
[24] Shaoquan Wu, Jiwu Huang, , Daren Huang, and Yun Q. Shi, "Efficiently Self – Synchronized Audio Watermarking for Assured Audio Data Transmission", *IEEE Transaction on Broad Casting*, vol. 5, no 1, pp. 69-76, March 2005.
[25] Pholsomboon S. and Vongpradhip S., "Rotation, scale, and translation resilient digital watermarking based on complex exponential function", *TENCON 2004. IEEE Region 10 Conference***,** page(s): 307-310 Vol. 1**,** 2004.