# Modified Patchwork Algorithm: A Novel Audio Watermarking Scheme

In-Kwon Yeo and Hyoung Joong Kim, *Member, IEEE*

***Abstract*—This paper presents the Modified Patchwork Algorithm (MPA), a statistical technique for audio watermarking algorithm in the transform (not only Discrete Cosine Transform (DCT), but also DFT and DWT) domain. The MPA is an enhanced version of the conventional patchwork algorithm. The MPA is sufficiently robust to withstand some attacks defined by Secure Digital Music Initiative (SDMI). Experimental results show that the proposed watermarking algorithm is sustainable against compression algorithms such as MP3 and AAC, as well as common signal processing manipulation attacks.**

## I. Introduction

MANY watermarking algorithms are now available [1]–[6], [8], [9], [11], [13]–[16]. Most of them are applicable to image or video. However, audio watermarking algorithms are comparatively few in number [1], [3], [4], [8], [11], [13], [14], [16]. Audio watermarking algorithms are not easy to develop due to several reasons. Most of all, the human ear is far more sensitive than other human sensory organs like eyes. Human ears can detect even a small amount of embedded noise especially when the signal power is weak.

It is believed that audio watermark is a good deterrent to illicit copying and dissemination of copyrighted audio. It can provide evidence of copyright infringements after the copyright violation has occurred. Audio watermark information is a special kind of bit pattern that is, in a strict sense, an intentional noise inserted into the digital audio to identify the copyright information such as author, recording label, and usage rules. Moreover, it can possibly be used to control the number of copies and even to trace the fingerprint of pirates. In addition, it can be used to resolve rightful ownership [7].

Audio watermarking algorithm must satisfy at least two constraints: inaudibility and robustness. Embedded audio watermarks should be almost inaudible. Also, the algorithm should be robust enough to withstand attempts such as removal or alteration of inserted watermarks. These two constraints may seem to be contradictory. However, they must be satisfied.

Patchwork [3] is a refreshingly novel algorithm. The algorithm was proposed for image watermarking. The authors inserted information into the time-domain data. However, it needed considerable effort before it could be fully implemented. This algorithm was not successful in audio watermarking applications. Nonetheless, it provides us with a solid base as an excellent tool for information hiding. Recently, Arnold [1] tried to improve the performance of the original patchwork algorithm. He proposed a nice idea that enhances the effectiveness of watermarks. He applied the algorithm to audio, especially in the frequency-domain. However, Arnold's algorithm also needs to be refined further. Moreover, performance of the algorithm was not very encouraging.

Anyhow, this algorithm [1] is a landmark in the area of watermark research. Our effort is focused on improving these previous patchwork algorithms. We also derive mathematical formulations that help to improve robustness. Of course, these mathematical formulations themselves are not sufficient to defend all tricky attacks such as time scale modifications and linear speed changes [12]. Multiple algorithms should be combined to defeat multiple attacks. However, our objective is to convey the core idea of the modified patchwork algorithm (MPA) [16], which can enhance the power of the original patchwork algorithm considerably. In Section II, previous work on the patchwork algorithms are reviewed and problems requiring elaborate treatment are identified. In Sections III and IV, embedding and detection techniques of MPA are presented based on mathematical derivations. Performance results of the pure MPA are provided along with the MPA with some modifications in Section V. Section VI concludes the paper.

## II. Review of the Patchwork Algorithm

Patchwork is an excellent watermarking algorithm proposed for images [3]. Bender *et al.* proposed the core idea. This algorithm embeds a special statistic into a host image. The two major steps in the algorithm are: (i) choose two patches pseudo-randomly and (ii) add the small constant value $d$ to the sample values of one patch $A$ and subtract the same value $d$ from the sample values of another patch $B$. Mathematically speaking

$$a_i^* = a_i + d, \quad b_i^* = b_i - d$$

where $a_i$ and $b_i$ are sample values of the patchwork sets $A$ and $B$, respectively. Thus, the original sample values have to be slightly modified. The detection process starts with the subtraction of the sample values between two patches. Then, $E[\bar{a}^* - \bar{b}^*]$, the expected value of the differences of the sample means is used to decide whether the samples contain watermark information or not, where $\bar{a}^*$ and $\bar{b}^*$ are sample means of the individual sample

Manuscript received September 12, 2001; revised July 11, 2002. This work was supported in part by the Institute of Information Technology Assessment, Korea, under Grant 0700004-1-1 (2000247) and Advanced Information Technology Research Center (AITrc). The associate editor coordinating the review of this manuscript and approving it for publication was Prof. C.-C. Jay Kuo.

I.-K. Yeo is with the Division of Mathematics and Statistical Informatics, Chonbuk National University, Jeonju 561-756, Korea (e-mail: inkwon@moak.chonbuk.ac.kr).

H. J. Kim is with the Department of Control and Instrumentation Engineering, Kangwon National University, Chunchon 200-701, Korea (e-mail: khj@kangwon.ac.kr).

Digital Object Identifier 10.1109/TSA.2003.812145

$a_i^*$ and $b_i^*$, respectively. Since two patches are used rather than one, it can detect the embedded watermarks without the original host images, which makes it a blind watermarking algorithm.

Patchwork itself is a very good algorithm, but it has some inherent drawbacks. Note that

$$E[\bar{a}^* - \bar{b}^*] = E[(\bar{a} + d) - (\bar{b} - d)] = E[\bar{a} - \bar{b}] + 2d$$

where $\bar{a}$ and $\bar{b}$ are sample means of the individual sample $a_i$ and $b_i$, respectively. The patchwork algorithm assumes that $E[\bar{a}^* - \bar{b}^*] = 2d$ due to the prior assumption that random sample ensures that expected values are all the same such that $E[\bar{a} - \bar{b}] = 0$. However, the actual difference of sample means, $\bar{a} - \bar{b}$, is not always zero in practice. Although the watermarked distribution is shifted to the right as shown in Fig. 1, the probability of a wrong detection still remains, the area smaller than 0 in the watermarked distribution. The performance of the patchwork algorithm depends on the distance between two sample means and $d$ which affects inaudibility. Furthermore, the patchwork algorithm has originally been designed for images. However, image is quite different from audio in many respects. Thus, it is impractical to apply this algorithm directly to audio watermarking.

The original patchwork algorithm has been applied to the spatial-domain (or, equivalently, time-domain in audio) data. However, time-domain embedding is vulnerable even to weak attacks and modifications. In fact, Arnold [1] has applied the patchwork algorithm in the frequency-domain, but not in the time-domain. Another important contribution of his work was the introduction of the power density function to the hypothesis tests. In addition, his algorithm is different from the original patchwork algorithms in several points.

1) Mean and variance of the sample values are computed in order to detect the watermarks. In the original algorithm [3], only mean was used.
2) Arnold [1] assumes that the distribution of the sample values is normal. The original algorithm [3] assumes that the distribution is uniform.
3) Embedding function of Arnold is multiplicative such as $a_i^* = a_i(1+d)$ and $b_i^* = b_i(1-d)$. The original algorithm [3] was based on the additive embedded function.
4) Arnold [1] tries to decide the value $d$ adaptively. However, the original algorithm [3] selected a fixed value 0.15 in the experiment.

The embedding function gives rise to considerable changes in scale when $d$ is large since it is multiplicative. Thus, serious quality degradation is expected. The expected value of the sample means is given as follows:

$$\begin{aligned} E[\bar{a}^* - \bar{b}^*] =& E[\bar{a}(1+d) - \bar{b}(1-d)] \\ =& E[\bar{a} - \bar{b}] + dE[\bar{a} + \bar{b}]. \end{aligned}$$

If $E[\bar{a} + \bar{b}] = 0$, then $E[\bar{a}^* - \bar{b}^*] = 0$ and, consequently, it cannot contribute to widen the gap between the two peaks shown in Fig. 1. It means that the peaks are too closely placed to detect watermarks. In order to widen the gap the value must be large. Unfortunately, it makes the audio quality worse. Thus, it is required that $E[\bar{a} + \bar{b}] \neq 0$ for successful application. However, sometimes this requirement may not hold, for instance in the DCT domain. Another contribution of Arnold [1] is the idea of deciding the value $d$ adaptively. This is a very important concept since it can allow watermarking algorithm to hide information according to the characteristics of the audio. Then, inaudibility and robustness can be enhanced considerably.

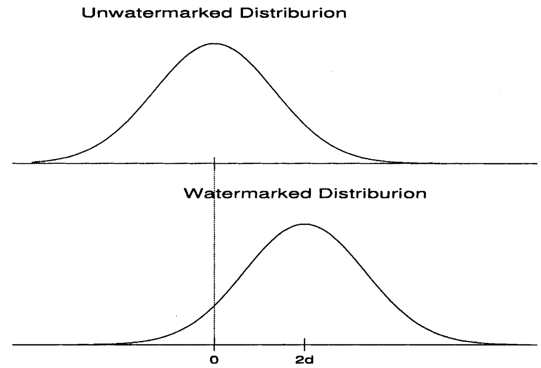



Fig. 1. Comparison of the unwatermarked and watermarked distributions of the mean difference.

## III. Embedding the Watermarks

The proposed algorithm in this paper inserts watermarks in the frequency domain. Even though the proposed algorithm is based on the DCT domain, it can be applied to all frequency domains such as DFT and DWT. Let $N$ be the size of a block in which the DCT is applied and one bit code is embedded. In order to make the statistical decision-making procedure effective, the bit code is repeatedly and consecutively embedded $R$ times. Repeated embedding of the same information $R$ times and detection based on majority voting plays error correcting functionality.

Embedding steps are summarized as follows.

1) Map the secret key and the watermark to the seed of a random number generator. Next, generate an index set $I = \{I_1, \ldots, I_{2n}\}$ whose elements are pseudo-randomly selected integer values from $[K_1, K_2]$, $1 \leq K_1 < K_2 \leq N$. Note that two index sets, $I^0$ and $I^1$, are needed to denote 1 bit code, 0 and 1, respectively. Distinct multiple index sets can also be used to designate multiple bits of code information in just one block. For instance, we can express 2 bits of code information with four different index sets. The choice of $K_1$ and $K_2$ is a crucial step in embedding the watermark because these values control the robustness and the inaudibility of the watermark.
2) Let $F = \{F_1, \ldots, F_N\}$ be the DCT coefficients whose subscript denote frequency range from the lowest to the highest frequencies. Define $A = \{a_1, \ldots, a_n\}$ as the subset of $F$ whose subscript corresponds to the first $n$ elements of the index set $I^0$ or $I^1$ according to the embedded code with a similar definition for $B = \{b_1, \ldots, b_n\}$ with the last $n$ elements, that is, $a_i = F_{I_i}$ and $b_i = F_{I_{n+i}}$, for $i = 1, \ldots, n$.

3) Calculate the sample means $\bar{a} = n^{-1}\sum_{i=1}^{n} a_i$ and $\bar{b} = n^{-1}\sum_{i=1}^{n} b_i$, respectively, and the pooled sample standard error

$$S = \sqrt{\frac{\sum_{i=1}^{n}(a_i-\bar{a})^2 + \sum_{i=1}^{n}(b_i-\bar{b})^2}{n(n-1)}}.$$

4) The embedding function presented below introduces a location-shift change

$$\begin{aligned} a_i^* &= a_i + \mathrm{sign}(\bar{a}-\bar{b})\sqrt{C}\frac{S}{2}, \\ b_i^* &= b_i - \mathrm{sign}(\bar{a}-\bar{b})\sqrt{C}\frac{S}{2} \end{aligned} \tag{1}$$

where $C$ is a constant and "sign" is the sign function. This function makes the large value set larger and the small value set smaller so that the distance between two sample means is always bigger than $d = \sqrt{C}S$ as shown in Fig. 2.

5) Finally, replace the selected elements $a_i$ and $b_i$ by $a_i^*$ and $b_i^*$, respectively, and then apply the inverse DCT.

Sometimes, it may so happen that the status about copyright information must be modified. Thus, the watermark has to be "re-marked". For example, every time when the user makes a copy, re-marking should be done in order to decrease the allowable number of copies. The copy control information (CCI) of the Secure Digital Music Initiative (SDMI) specification [11] allows a 2-bit code to represent the copy permission status. For instance, the code "10" stands for "one-copy-only," and "11" stands for "no-more-copy." After copying a watermarked file, the copy permission status should be changed from "one-copy-only" to "no-more-copy." That is, the second bit must be re-marked from "0" to "1." For this kind of task too, we present the re-marking procedure. However, the MPA requires a procedure similar to the remarking step.

Suppose that code "0" should be replaced by "1" after re-marking. To do this, obtain subsets $A_0 = \{a_{01}^*, \ldots, a_{0n}^*\}$ and $B_0 = \{b_{01}^*, \ldots, b_{0n}^*\}$ from the index set $I^0$, $A_1 = \{a_{11}, \ldots, a_{1n}\}$ and $B_1 = \{b_{11}, \ldots, b_{1n}\}$ from the index set $I^1$, all from $F = \{F_1, \ldots, F_N\}$ and compute the sample means $\bar{a}_0^*$, $\bar{b}_0^*$, $\bar{a}_1$, and $\bar{b}_1$ and the pooled sample standard errors $S_0^*$ and $S_1$. The cleaning function is given as follows:

$$\begin{aligned} a_{0i} &= a_{0i}^* - \mathrm{sign}(\bar{a}_0^*-\bar{b}_0^*)\sqrt{C}\frac{S_0^*}{2}, \\ b_{0i} &= b_{0i}^* + \mathrm{sign}(\bar{a}_0^*-\bar{b}_0^*)\sqrt{C}\frac{S_0^*}{2}. \end{aligned}$$

Replace the marked coefficients by the cleaned coefficients. Next, the embedding function is again applied to the cleaned coefficients. The embedding function is restated as follows:

$$\begin{aligned} a_{1i}^* &= a_{1i} + \mathrm{sign}(\bar{a}_1-\bar{b}_1)\sqrt{C}\frac{S_1}{2}, \\ b_{1i}^* &= b_{1i} - \mathrm{sign}(\bar{a}_1-\bar{b}_1)\sqrt{C}\frac{S_1}{2}. \end{aligned}$$

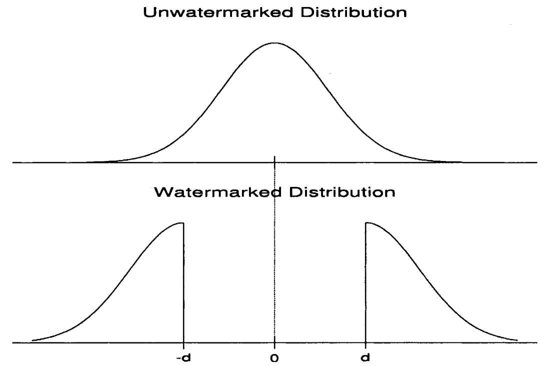



Fig. 2. A comparison of the un-watermarked and watermarked distributions of the mean difference by the MPA.

## IV. Detecting the Watermark and Test Statistic

Since the proposed embedding function (1) introduces relative changes of two sets in location, a natural test statistic which is used to decide whether or not the watermark is embedded should concern the distance between the means of $A$ and $B$. In this section, we present the detecting algorithm and investigate statistical properties of the test statistic. The decoding process is as follows.

1) Map the secret key and watermark to the seed of random number generator and then generate the index sets $I^0$ and $I^1$ which was applied to the encoding process.
2) Obtain the subsets $A_1$ and $B_1$ from $F = \{F_1, \ldots, F_N\}$ and compute the sample means and the pooled sample standard errors. Obtain the subsets $A_0 = \{a_{01}, \ldots, a_{0n}\}$ and $B_0 = \{b_{01}, \ldots, b_{0n}\}$ from the index set $I^0$, $A_1 = \{a_{11}, \ldots, a_{1n}\}$ and $B_1 = \{b_{11}, \ldots, b_{1n}\}$ from the index set $I^1$, all from $F = \{F_1, \ldots, F_N\}$ and compute the sample means $\bar{a}_0$, $\bar{b}_0$, $\bar{a}_1$, and $\bar{b}_1$ and the pooled sample standard errors $S_0$ and $S_1$.
3) Calculate the test statistics

$$T_0^2 = \frac{(\bar{a}_0-\bar{b}_0)^2}{S_0^2} \text{ and } T_1^2 = \frac{(\bar{a}_1-\bar{b}_1)^2}{S_1^2}$$

and define $T^2$ as the larger value obtained from two statistics.
4) Compare $T^2$ with the threshold $M$ and decide that watermark is embedded if $T^2 > M$. Only when $T^2 > M$, 0 is assigned if $T_0^2 > T_1^2$, and 1 otherwise.

In order to investigate the statistical properties of our test data, we first set the null hypothesis $(H_0)$ and alternative hypothesis $(H_1)$ as follows:

$H_0$: audio signal is not watermarked;
$H_1$: audio signal is watermarked.

Then, two kinds of false detection are incorporated in hypothesis testing:

Type I error : Rejection of $H_0$ when $H_0$ is true;
Type II error: Non-rejection of $H_0$ when $H_1$ is true.

A good watermark algorithm might reduce the possibility of both errors as much as possible.

Next, the probabilities of Type I error and Type II error occurring are studied to show the effectiveness of our method from the statistical point of view. Let $\mu$ be the mean of $K_2 - K_1 + 1$

DCT coefficients on frequency domain $[K_1, K_2]$ and let $\sigma^2$ be the variance. Assume that the DCT coefficients are normally distributed. Then, under $H_0$ and random sampling scheme

$$T^2 = \frac{(\bar{a} - \bar{b})^2}{S^2} \sim F_{1,2n-2}$$

where $F_{\nu_1,\nu_2}$ denotes $F$-distribution with $\nu_1$ and $\nu_2$ degrees of freedom. Let $p$ be the probability of the event that $T_0^2 > M$ when the watermark is not embedded. Then, the probability of Type I error, $\alpha$, is calculated as

$$\alpha = P\left(T_0^2 > M \cup T_1^2 > M|H_0\right) = p(2-p).$$

Since we repeatedly mark one bit code $R$ times, we have to test the bit code $R$ times to extract the information. Let $X$ be the number of events that $T_0^2$ or $T_1^2$ is greater than $M$ out of $R$ tests. Suppose that we finally conclude that, for a constant $Z$, a watermark is marked when $X \geq Z$. Then, the probability of Type I error for the final conclusion can be computed by using the binomial probability density function

$$\begin{aligned}\bar{\alpha} =& P(\text{Type I error}) = P(X \geq Z) \\ =& \sum_{x=Z}^{R} \binom{R}{x} \alpha^x (1-\alpha)^{R-x}.\end{aligned}$$

Next, we investigate the probabilities of Type II error, $\beta$ and $\bar{\beta}$, where $\beta$ and $\bar{\beta}$ are defined as

$$\begin{aligned}\beta =& P\left(T_0^{2*} \leq M \cap T_1^{2*} \leq M|H_1\right), \\ \bar{\beta} =& P(X < Z) = \sum_{x=0}^{Z-1} \binom{R}{x} \beta^x (1-\beta)^{R-x}\end{aligned}$$

respectively. Since the embedding function shifts the location of the two sets, there may be no effect on the variance if the samples are not overlapped. So $S' \simeq S$ where $S'$ stands for the pooled standard error of watermarked samples. Let $T$ be the square root version of the proposed test statistic, that is, $T = (\bar{a} - \bar{b})/S$. Then, the proposed test statistics under $H_1$ can be written as

$$T^{2^*} \simeq (T + \text{sign}(T)\sqrt{C})^2 \geq T^2 + C.$$

Setting $C \geq M$, we can make the test statistic always greater than the threshold $M$ as shown in Fig. 3. This implies that the proposed algorithm can make probability for Type II error zero when there is no malicious attack.

Finally, we consider another type of the false detection that "0" is marked but "1" is detected or vice versa. The probability of this type of false decision is called the bit error rate (BER). Under normality assumption, the probability of false decision that "0" is marked but "1" is detected can be obtained by numerically calculating

$$\begin{aligned}\gamma =& P(T_1^2 \geq T_0^2, T_1^2 > M|0 \text{ is embeded}) \\ =& \int_0^\infty \int_C^{t_1} \phi(t_1 > M) f_0(t_0) f_1(t_1) dt_0 dt_1\end{aligned}$$

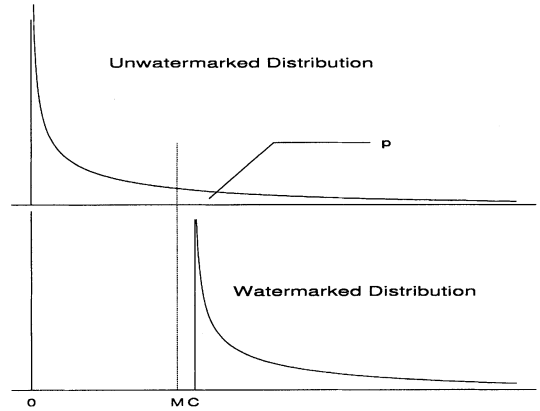



Fig. 3. Comparison of the un-watermarked and watermarked distributions of the detection function.

where

$$f_0(t) = \frac{\left(\frac{1}{2n-2}\right)^{1/2} \Gamma\left(\frac{2n-1}{2}\right)^2 (t-C)^{-1/2}}{\Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{2n-2}{2}\right)^2 \left(1 + \frac{t-C}{2n-2}\right)^{(2n-1)/2}},$$

$$f_1(t) = \frac{\left(\frac{1}{2n-2}\right)^{1/2} \Gamma\left(\frac{2n-1}{2}\right)^2 t^{-1/2}}{\Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{2n-2}{2}\right)^2 \left(1 + \frac{t}{2n-2}\right)^{(2n-1)/2}}$$

and $\phi(t > M) = 1$ if $t > M$, 0 otherwise. If we come to a conclusion after testing $R$ times, then the final BER will be

$$\bar{\gamma} = P(\text{BER}) = P(X \geq Z) = \sum_{x=Z}^{R} \binom{R}{x} \gamma^x (1-\gamma)^{R-x}.$$

## V. Robustness and Inaudibility Tests

To test the robustness of the proposed watermarking algorithm, a bit code was repeatedly marked 10 times and a total of 9050 blocks were tested to determine the error rate. All audio signals to be tested were 16 bits signed stereo sampled at 44.1 kHz. The audio samples include Rock, Jazz, and Classical music. The size of the block was set as $N = 4410$ for the original signal and $N = 2205$ for down-sampled signal to 22.05kHz. We generated 50 random numbers $(n = 25)$ for each index set whose elements were randomly chosen from $[K_1, K_2] = [500, 1000]$ paying regard to imperceptibility. The location-shift change parameter $C = 12$ and the threshold $M = 4$ were adopted. We assume that the signal is watermarked when $X \geq 5$ in the repetition scheme. Then, the theoretical probabilities of Type I error are $\alpha = 0.051$ and $\bar{\alpha} = 6.7 \times 10^{-5}$, respectively. The observed probabilities of Type II error were $\alpha = 0.082$ and $\bar{\alpha} = 0.000$, respectively. To test the robustness of the algorithm against various types of manipulations, the following types of signal processing were employed.

1) Down-sampling: The watermarked audio signals with a sampling rate of 44.1 kHz were reduced to 22.5 kHZ.
2) Band-pass filtering: The cut-off frequencies were 100 Hz for the low-pass and 6 kHz for the high-pass filter. The second order Butterworth filter was applied.

3) Echo addition: The echoed signals of which the maximum delay time is 100 ms were mixed in the watermarked signals.
4) Equalization: The watermarked signals were amplified at 10 band graphic equalizer with $+6$ dB or $-6$ dB gain.
5) Low bit-rate codec: The robustness against the low-rate codecs was tested by using MPEG 1 Layer III compression and AAC with compression rates of 64, 96, and 128 kbps. Since the unexpected number of zeros are added at the front of audio signal during decompression, a preliminary work to obtain the information about the start position of watermarking was done.

The experimental results are given in Table I which shows that the watermark is not affected by down-sampling, equalization, and high bit rate compressions and the major portions still survive at low bit-rate compression and against band-pass filtering, especially when the repetition scheme is adopted. Good error-correcting codes can enhance detection rate further. However, they are beyond the scope of our paper since, as is mentioned in Section I, our objective is to convey the core idea of the modified patchwork algorithm (MPA) [16]. In order to design error-correcting codes we have to compute the bit error rates, which depend not only on the watermarking algorithm per-se but also on the attacks. An experiment shows that an adaptive scheme for generating random numbers can improve the robustness without losing the quality of audio signal.

For audio signal with high energy, the watermark tends to be imperceptible even though the low frequency coefficients are modified. Since the relatively low frequency components are perceptually significant, not much data loss occurs at these regions and so the reliability of watermark may be enhanced.

Existing methods [4], [14] compute psycho-acoustic thresholds and embed spread-spectrum noise below the threshold in order not to be audible. However, the psycho-acoustic model can be applied only when DFT is used. Note that the psycho-acoustic model does not exist for DCT and DWT. We do not consider the psycho-acoustic model since our method can be applied to any transform domain, DCT and DWT as well as DFT. Thorough experiments show that the region around $[K_1, K_2] = [500, 1000]$ is better for both robustness and inaudibility. Of course, the best regions depend on the music. Inaudibility is a quite subjective measure. Anyhow, we can hardly hear the noise or audible artifacts from the watermarked audio.

TABLE I
ERROR PROBABILITIES FOR DIFFERENT MANIPULATIONS

| | Type II (%) | | BER (%) | |
|---|---|---|---|---|
| Manipulation | $\beta$ | $\bar{\beta}$ | $\gamma$ | $\bar{\gamma}$ |
| No manipulation | 0.0 | 0.0 | 0.0 | 0.0 |
| Down-Sampling | 0.0 | 0.0 | 0.0 | 0.0 |
| Band-pass filtering | 20.3 | 1.5 | 1.4 | 0.0 |
| Echo addition | 12.0 | 0.1 | 0.9 | 0.0 |
| Equalization | 0.0 | 0.0 | 0.0 | 0.0 |
| MPEC 1 Layer III (128) | 0.2 | 0.0 | 0.1 | 0.0 |
| MPEC 1 Layer III (96) | 1.0 | 0.0 | 0.2 | 0.0 |
| MPEC 1 Layer III (64) | 2.1 | 0.0 | 0.2 | 0.0 |
| AAC (128) | 0.0 | 0.0 | 0.0 | 0.0 |
| AAC (96) | 1.6 | 0.1 | 0.2 | 0.0 |
| AAC (64) | 17.1 | 4.0 | 0.8 | 0.0 |

## VI. Concluding Remarks

This paper presents a core idea of MPA (Modified Patchwork Algorithm). Main contributions of this paper include the following.

1) The embedding factor $d$ is calculated adaptively based on the sample mean and sample variance. Adaptive patchwork algorithm is very important in two respects. It can maintain inaudibility while keeping sufficient robustness. In addition, adaptive patchwork algorithm is more robust against specific attacks such as copy attack [10] since its value $d$ is computed based on the host signals.
2) Sign function in embedding functions is used. It enhances the detection rate and consequently reduces false positive and false negative errors. The sign function itself is simple, but very difficult to derive mathematical formula since it is a nonlinear function.
3) Patch size is around 50 samples in transform domain. Even with such small samples we can achieve acceptable robustness. Since the sample size is so small, inaudibility characteristics are good. Of course, the 50 samples are spread over the whole samples, in our case, over 4410 samples when we transform them inversely to the time domain. In this respect, the patchwork algorithm is considered to be a kind of spread-spectrum algorithm.

The three facts make the MPA algorithm very robust. Simulation results show that its performance improves considerably over existing patchwork algorithms. Moreover, its inaudibility characteristics are good. However, the MPA itself is not a cure-for-all and should be incorporated with other watermarking algorithms. The MPA itself still needs refinements. For example, it can be incorporated with psycho-acoustic model when the DFT is used for inaudibility and error correcting codes for more robustness.

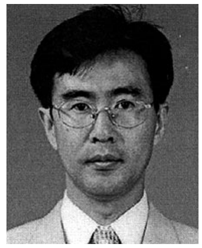
## Acknowledgment

The authors appreciate Prof. D. Ghose of Indian Institute of Science for his comments and proofreading.

## References

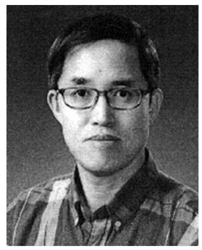
[1] M. Arnold, "Audio watermarking: Features, applications and algorithms," in *IEEE Int. Conf. Multimedia and Expo 2000*, vol. 2, 2000, pp. 1013–1016.
[2] M. Barni, F. Bartolini, V. Cappellini, and A. Piva, "DCT-domain system for robust image watermarking," *Signal Process.*, vol. 66, pp. 357–372, 1998.

[3] W. Bender, D. Gruhl, N. Morimoto, and A. Lu, "Techniques for data hiding," *IBM Syst. J.*, vol. 35, no. 3/4, pp. 313–336, 1996.
[4] L. Boney, A. H. Tewfik, and K. N. Hamdy, "Digital watermarking for audio signals," in *Proc. 3rd Int. Conf. Multimedia Computing and Systems*, 1996, pp. 473–480.
[5] I. J. Cox, M. L. Miller, and J. A. Bloom, *Digital Watermarking*. San Mateo, CA: Morgan Kaufmann, 2002.
[6] I. Cox, J. Kilian, F. T. Leighton, and T. Shamoon, "Secure spread spectrum watermarking for multimedia," *IEEE Trans. Image Processing*, vol. 6, pp. 1673–1687, 1997.
[7] S. Craver, N. Memon, B.-L. Yeo, and M. M. Yeung, "Resolving rightful owenerships with invisible watermarking techniques: Limitations, attacks, and implications," *IEEE J. Select. Areas Commun.*, vol. 16, pp. 573–586, 1998.
[8] J. Dittmann, M. Steinebach, and R. Steinmetz, "Digital watermarking for MPEG audio layer 2," in *Proc. ACM Multimedia*, 1999, pp. 117–122.
[9] F. Hartung and B. Girod, "Watermarking of uncompressed and compressed video," *Signal Process.*, vol. 66, no. 3, pp. 283–301, 1998.
[10] M. Kutter, S. Volshynovskiy, and A. Herrigel, "The watermark copy attack," in *Security and Watermarking of Multimedia Contents II*, 2000, vol. 3971, pp. 226–239.
[11] L. Qiao and K. Nahrstedt, "Non-invertible watermarking methods for MPEG encoded audio," in *IS & T/SPIE Conf. Security and Watermarking of Multimedia Contents*, 1999, pp. 194–202.
[12] "Call for Proposal," SDMI, http://www.sdmi.org/cfp.html, 2000.
[13] J. W. Seok and J. W. Hong, "Audio watermarking for copyright protection of digital audio data," *Electron. Lett.*, vol. 37, no. 1, pp. 60–61, 2001.
[14] M. D. Swanson, B. Zhu, A. H. Tewfik, and L. Boney, "Robust audio watermarking using perceptual coding," *Signal Process.*, vol. 66, no. 3, pp. 337–356, 1998.
[15] R. B. Wolfgang, C. I. Podilchuk, and E. J. Delp, "Perceptual watermarks for digital images and video," *Proc. IEEE*, vol. 87, pp. 1108–1126, 1999.
[16] I. K. Yeo and H. J. Kim, "Modified patchwork algorithm: The novel audio watermarking scheme," in *Proc. IEEE Int. Conf. Information Technology: Coding and Computing*, Las Vegas, NV, April 2–4, 2001, pp. 237–242.



**In-Kwon Yeo** received the Ph.D. degree in statistics from University of Wisconsin-Madison in 1997.

He joined the Department of Control and Instrumentation Engineering, Kangwon National University, as a Visiting Professor in 2000. He is currently an Assistant Professor at the Division of Mathematics and Statistical Informatics, Chonbuk National University, Jeonju, Korea. His current research interests include transformations, pattern classifications of time series data, and bio-informatics.



**Hyoung Joong Kim** (S'86–M'89) received the B.S., M.S., and Ph.D. degrees from Seoul National University, Seoul, Korea, in 1978, 1986, 1989, respectively.

He joined the faculty of the Department of Control and Instrumentation Engineering, Kangwon National University, Chunchon, Korea, in 1989, where he is currently a Professor. He was a Visiting Scholar at the University of Southern California, Los Angeles, during 1992–1993. During 1998–2000, he was the Prime Investigator of the iPCTV Project developing interactive digital television, and implemented MHP, ATVEF, and DASE specifications and performed conformance tests. He is the Prime Investigator of the iMS (Interactive Media Solution) developing the end-to-end solution for data broadcasting system. He was the founder of the International Workshop on Digital Watermarking (IWDW) and is a co-chair of the Technical Program Committee of the IWDW2002. Recently, his research team contributed Digital Item Adaptation technologies for MPEG-21. His research interests include parallel computing, multimedia computing, and multimedia security.

Dr. Kim served as a Guest Editor for the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY.