# Robust Patchwork-Based Embedding and Decoding Scheme for Digital Audio Watermarking

Iynkaran Natgunanathan, Yong Xiang, Yue Rong, *Senior Member, IEEE*, Wanlei Zhou, *Senior Member, IEEE*, and Song Guo, *Senior Member, IEEE*

*Abstract*—This paper presents a novel patchwork-based embedding and decoding scheme for digital audio watermarking. At the embedding stage, an audio segment is divided into two subsegments and the discrete cosine transform (DCT) coefficients of the subsegments are computed. The DCT coefficients related to a specified frequency region are then partitioned into a number of frame pairs. The DCT frame pairs suitable for watermark embedding are chosen by a selection criterion and watermarks are embedded into the selected DCT frame pairs by modifying their coefficients, controlled by a secret key. The modifications are conducted in such a way that the selection criterion used at the embedding stage can be applied at the decoding stage to identify the watermarked DCT frame pairs. At the decoding stage, the secret key is utilized to extract watermarks from the watermarked DCT frame pairs. Compared with existing patchwork watermarking methods, the proposed scheme does not require information of which frame pairs of the watermarked audio signal enclose watermarks and is more robust to conventional attacks.

*Index Terms*—Audio watermarking, discrete cosine transform (DCT), generalized Gaussian distribution, patchwork.

## I. INTRODUCTION

WITH the advances of multimedia and Internet technologies, digital data can be easily reproduced, manipulated and distributed without any quality degradation. This has resulted in strong demand for preventing illegal use of copyrighted data. Digital watermarking is an important technique for copyright protection and integrity authentication in an open network environment [1]–[3]. Technically speaking, digital watermarking aims to hide watermark data (such as publisher's name, signature, logo, ID number, etc.) into the actual media object without affecting its normal usage. When necessary, the owners can extract the watermark data to declare their copyright [4]–[6]. Based on the application areas, digital watermarking is usually categorized into audio watermarking, image watermarking and video watermarking [4], [5], [7]–[9]. This paper limits its attention to audio watermarking. Since an audio signal is one-dimensional and the human auditory perception is more sensitive than other sensory perceptions such as vision [5], [6], [10], it is more difficult to hide additional information into an audio signal than into other multimedia data, without lowering the quality of the media object.

An effective and practical audio watermarking scheme should exhibit three important characteristics: imperceptibility, robustness and security. Imperceptibility denotes that the embedded watermark data should be almost inaudible. Robustness refers to the ability of recovering the watermark data from the watermarked signal in the absence and presence of attacks. The requirements on imperceptibility and robustness are contradictory but must be satisfied. Security means that a secret key should be used in the watermarking scheme such that an unauthorized person cannot extract the watermarks without knowing the secret key. Apart from these attributes, low computational complexity and adjustability of the watermarking scheme are additional advantages. An efficient watermarking scheme is particularly important for time-demanding applications (e.g., delivering the audio data over the Internet), and an adjustable watermarking scheme makes itself suitable for various applications. Furthermore, it is desirable that the embedded watermarks can be extracted at the decoding stage without resort to the host audio signal [3].

In recent years, many watermarking methods have been proposed for audio signals. These audio watermarking methods use different techniques such as spread spectrum [11]–[13], echo-hiding [5], [6], [14]–[18], support vector regression [1], [19], [20], and patchwork [2], [3], [10], [21]. Among existing watermarking methods, the patchwork-based methods show great potential to resisting conventional attacks such as amplification, re-quantization, re-sampling, noise addition, and lossy compression [e.g., MP3 and advanced audio coding (AAC)] attacks. They can also achieve good imperceptibility and high level of security.

Patchwork watermarking technique was initially developed for images [22]. Then Arnold extended this technique to audio signals [10], followed by a modified patchwork algorithm [2] which was proposed by Yeo and Kim to improve watermarking performance. The patchwork methods in [2] and [10] utilize, respectively, the discrete cosine transform (DCT) coefficients and

Fourier transform coefficients of a host audio segment to form four subsets, called patches. The digital watermark "0" or "1" is embedded into the audio segment by modifying the patches according to certain embedding rule. The performance of the methods in [2] and [10] relies on the assumption that the chosen patches have the same statistical property. This assumption is not always true in practice due to the finite length of patches. In [3], Kalantari *et al.* proposed a multiplicative patchwork method to solve this problem, where two patches are produced for each host audio segment by using its wavelet transform coefficients. If the two patches of an audio segment have comparable statistical characteristics, the audio segment is used to embed watermark. Otherwise, the audio segment is excluded from inserting watermark. In this method, a sizeable percentage of audio segments are considered to be unsuitable for watermark embedding. Since watermarks are only embedded into some segments of the host audio signal, it is important to know, at the decoding end, which segments of the watermarked signal contain watermarks. Without this information, considerable false watermarks will be "extracted" from the unwatermarked audio segments. However, [3] does not provide a way to find the watermarked segments. It should be noted that the approach proposed to estimate the indices of selected image frames in [23] cannot be directly applied nor simply modified to identify the watermarked segments.

In this paper, we propose a patchwork-based watermarking scheme for audio signals. In the proposed scheme, the host audio segment is partitioned into two subsegments and then the DCT coefficients of the subsegments are calculated. After discarding those DCT coefficients corresponding to high frequency components, the remaining DCT coefficients are divided into multiple frame pairs. A criterion is used to select the DCT frame pairs that are appropriate for embedding watermarks. Under the control of a pseudonoise (PN) sequence, a watermark is embedded into the selected frame pairs by altering the associated DCT coefficients. The embedding algorithm is designed in such a way that the selection criterion utilized at the embedding stage can also be employed at the decoding stage to discover the watermarked frame pairs from the watermarked signal. After finding the frame pairs containing watermarks, one can extract the watermarks by using the PN sequence as a secret key. The new scheme is superior to the existing patchwork watermarking methods as it does not need any additional information to find the watermarked frame pairs at the decoding end and has higher robustness against conventional attacks. Its effectiveness is demonstrated by simulation results.

The remainder of the paper is organized as follows. The new patchwork-based embedding and decoding scheme is presented in Section II. Simulation results are provided in Section III to illustrate the performance of the proposed scheme. Section IV concludes the paper.

## II. PROPOSED SCHEME

### A. Watermark Embedding

*1) Generation of DCT Frames and Fragments:* Fig. 1 shows the process of generating DCT frames and fragments. First, the host audio signal is segmented to obtain a number of segments,
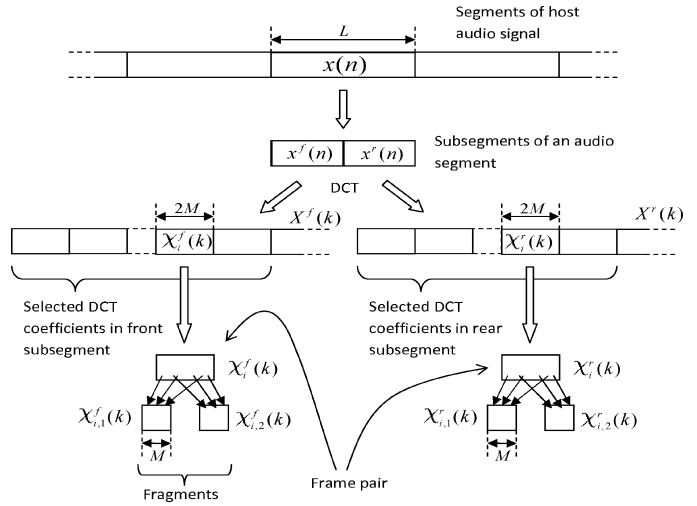


Fig. 1. Illustration of frame and fragment generation.

which are of equal length $L$, where $L$ is an even integer. If a host audio segment is suitable for watermark embedding, a digital watermark bit, which is either "1" or "0," will be inserted into the segment. Let $x(n)$ be a host audio segment. Then we further divide $x(n)$ into two equal-length subsegments, called front subsegment $x^f(n)$ and rear subsegment $x^r(n)$, respectively. Clearly, the length of $x^f(n)$ and $x^r(n)$ is $L/2$. Denote the DCTs of $x^f(n)$ and $x^r(n)$ by $X^f(k)$ and $X^r(k)$, respectively, which are defined as follows [24]:

$$X^f(k) = l(k) \sum_{n=0}^{L/2-1} x^f(n) \cos\left\{\frac{\pi(2n+1)k}{L}\right\} \quad (1)$$

$$X^r(k) = l(k) \sum_{n=0}^{L/2-1} x^r(n) \cos\left\{\frac{\pi(2n+1)k}{L}\right\} \quad (2)$$

where $k = 0, 1, \ldots, L/2 - 1$, and

$$l(k) = \begin{cases} \frac{1}{\sqrt{\frac{L}{2}}}, & \text{if } k = 0 \\ \sqrt{\frac{2}{\frac{L}{2}}}, & \text{if } 1 \leq k < \frac{L}{2}. \end{cases}$$

From (1) and (2), it can be seen that the DCT coefficients $X^f(k)$ and $X^r(k)$ are real-valued and have direct relationship with the frequency components of $x^f(n)$ and $x^r(n)$, respectively. Since high frequency components are vulnerable to compression attacks, we only use the DCT coefficients corresponding to a low to middle frequency range $f < f_{\text{limit}}$ to embed watermark. Thus, high-frequency components are freed by this process from being selected as potential DCT frames for embedding watermark. We denote the parts of $X^f(k)$ and $X^r(k)$ which are related to the selected frequency range by $\chi^f(k)$ and $\chi^r(k)$, respectively. To further improve robustness, one watermark bit can be implanted into an audio segment multiple times, say $P$ times. For this purpose, we split $\chi^f(k)$ into $R$ frames of length $2M$ and denote the $i$th frame of $\chi^f(k)$ by $\chi_i^f(k)$, $i = 1, 2, \ldots, R$. Similarly, $\chi^r(k)$ is also split into $R$ frames of length $2M$ and its $i$th frame is denoted by $\chi_i^r(k)$. We refer $\chi_i^f(k)$ and $\chi_i^r(k)$ as a DCT frame pair.

Security is a primary concern for any watermarking method. To ensure that our watermarking scheme is of high level of

security, we use a PN sequence $p(n)$ of length $2M$ to sort the DCT coefficients in a frame into two fragments, each of which contains $M$ DCT coefficients. Here, the $2M$-length PN sequence $p(n)$ is generated by randomly taking all the integers in $[0, (2M-1)]$. For example, if $2M = 30$, a possible $p(n)$ could be $p(n) = \{3, 0, 29, 1, \ldots 26, 28, 7\}$. Denote $p(n) = \{p(0), p(1), \ldots, p(2M-1)\}$, and let $k_i$ be the starting index of the frames $\chi_i^f(k)$ and $\chi_i^r(k)$. Then the two fragments corresponding to $\chi_i^f(k)$ can be written as

$$\chi_{i,1}^f(k) = \{\chi_i^f(k_i + p(0)), \chi_i^f(k_i + p(1))$$
$$\ldots, \chi_i^f(k_i + p(M-1))\} \tag{3}$$
$$\chi_{i,2}^f(k) = \{\chi_i^f(k_i + p(M)), \chi_i^f(k_i + p(M+1))$$
$$\ldots, \chi_i^f(k_i + p(2M-1))\} \tag{4}$$

and the fragments corresponding to $\chi_i^r(k)$ are

$$\chi_{i,1}^r(k) = \{\chi_i^r(k_i + p(0)), \chi_i^r(k_i + p(1))$$
$$\ldots, \chi_i^r(k_i + p(M-1))\} \tag{5}$$
$$\chi_{i,2}^r(k) = \{\chi_i^r(k_i + p(M)), \chi_i^r(k_i + p(M+1))$$
$$\ldots, \chi_i^r(k_i + p(2M-1))\}. \tag{6}$$

*2) Selection of DCT Frame Pairs:* Among the frame pairs $\{\chi_i^f(k), \chi_i^r(k)\}$, $i = 1, 2, \ldots, R$, some of them may not be suitable for watermark embedding as inserting watermarks into these frame pairs could lower the perceptual quality of the watermarked signal to an unacceptable level. In order to select appropriate frame pairs, we define the means of the absolute-valued fragments and frames as follows:

$$p_{i,j} = E\left(\left|\chi_{i,j}^f(k)\right|\right) \tag{7}$$
$$q_{i,j} = E\left(\left|\chi_{i,j}^r(k)\right|\right) \tag{8}$$
$$p_i = E\left(\left|\chi_i^f(k)\right|\right) \tag{9}$$
$$q_i = E\left(\left|\chi_i^r(k)\right|\right) \tag{10}$$

where $|a|$ denotes the absolute value of $a$, $E(\cdot)$ stands for averaging operation, $i = 1, 2, \ldots, R$ and $j = 1, 2$. From (3)–(10), it is easy to see that

$$p_i = \frac{p_{i,1} + p_{i,2}}{2} \quad \text{and} \quad q_i = \frac{q_{i,1} + q_{i,2}}{2} \tag{11}$$

for $i = 1, 2, \ldots, R$. We also define

$$\tilde{p}_i = |p_{i,1} - p_{i,2}| - \alpha_1 p_i \tag{12}$$
$$\tilde{q}_i = |q_{i,1} - q_{i,2}| - \alpha_1 q_i \tag{13}$$

where $\alpha_1$ is a small positive constant satisfying $\alpha_1 < 2$ and $i = 1, 2, \ldots, R$.

As will be shown in the watermark embedding rule later, watermarks are inserted by changing $p_{i,j}$ and $q_{i,j}$ (j = 1,2) according to the values of $p_i$ and $q_i$, respectively. Clearly, the smaller the changes on $p_{i,j}$ and $q_{i,j}$ are, the better the perceptual quality is. This requires the condition that $p_{i,1}$ and $p_{i,2}$ are close to $p_i$, and $q_{i,1}$ and $q_{i,2}$ are close to $q_i$. In other words, the difference between $p_{i,1}$ and $p_{i,2}$ and that between $q_{i,1}$ and $q_{i,2}$ are small. This condition can be satisfied if

$$\tilde{p}_i \leq 0 \quad \text{and} \quad \tilde{q}_i \leq 0.$$

The above two inequalities form a base for choosing frame pairs for watermark embedding. In order to avoid embedding watermarks into silent periods, we propose the following criterion to select frame pairs suitable for hiding watermarks:

$$\tilde{p}_i \leq 0, \quad \tilde{q}_i \leq 0, \quad \text{and} \quad \min\{p_i, q_i\} \geq \theta_{th} \tag{14}$$

where $\theta_{th}$ is a threshold, which can be determined empirically. If (14) holds, a watermark will be inserted into the $i$th frame pair $\{\chi_i^f(k), \chi_i^r(k)\}$. Otherwise, no watermark will be embedded into this frame pair. Usually, only some frame pairs are suitable for inserting watermark in practical applications. If none of the frame pairs in the subsegments $\{\chi^f(k), \chi^r(k)\}$ satisfies (14), the corresponding host audio segment will not be used for watermark embedding.

Obviously, the value of $\alpha_1$ has impact on imperceptibility and robustness. With the increase of $\alpha_1$, more frame pairs will be selected for watermark embedding, which leads to higher robustness. On the other hand, increasing $\alpha_1$ results in that $p_{i,j}$ and $q_{i,j}$ can differ more from $p_i$ and $q_i$, respectively, where $j = 1, 2$. Consequently, larger modifications will be made on the DCT coefficients to embed the watermark bit, which reduces perceptual quality.

*3) Watermark Embedding Rule:* Assume that the $l$th frame pair $\{\chi_l^f(k), \chi_l^r(k)\}$ is selected for watermark embedding. To insert a watermark bit into $\{\chi_l^f(k), \chi_l^r(k)\}$, we alter the means of the corresponding absolute-valued fragments $\left|\chi_{l,j}^f(k)\right|$ and $\left|\chi_{l,j}^r(k)\right|$, where $j = 1, 2$. Let $\chi_{l,j}^{\prime f}(k)$ and $\chi_{l,j}^{\prime r}(k)$ be the modified counterparts of $\chi_{l,j}^f(k)$ and $\chi_{l,j}^r(k)$, respectively. As shown in (7) and (8), the means of $\left|\chi_{l,j}^f(k)\right|$ and $\left|\chi_{l,j}^r(k)\right|$ are $p_{l,j}$ and $q_{l,j}$, respectively. Similarly, we denote the mean of $\left|\chi_{l,j}^{\prime f}(k)\right|$ by $p_{l,j}'$ and the mean of $\left|\chi_{l,j}^{\prime r}(k)\right|$ by $q_{l,j}'$. Let $\alpha_2$ be another small positive constant satisfying

$$0 < \alpha_2 < \alpha_1. \tag{15}$$

Based on these symbols, we propose the following watermark embedding rule.
- Embedding of watermark bit "0":
    *If* $(p_{l,1} - p_{l,2}) \geq \alpha_2 p_l$, *then*
    $p_{l,1}' = p_{l,1}$ *and* $p_{l,2}' = p_{l,2}$.
    *Otherwise,*
    $p_{l,1}' = (1 + 0.5 \times \alpha_2) p_l$ *and* $p_{l,2}' = (1 - 0.5 \times \alpha_2) p_l$.
    *If* $(q_{l,2} - q_{l,1}) \geq \alpha_2 q_l$, *then*
    $q_{l,1}' = q_{l,1}$ *and* $q_{l,2}' = q_{l,2}$.
    *Otherwise,*
    $q_{l,1}' = (1 - 0.5 \times \alpha_2) q_l$ *and* $q_{l,2}' = (1 + 0.5 \times \alpha_2) q_l$.
- Embedding of watermark bit "1":
    *If* $(p_{l,2} - p_{l,1}) \geq \alpha_2 p_l$, *then*
    $p_{l,1}' = p_{l,1}$ *and* $p_{l,2}' = p_{l,2}$.
    *Otherwise,*
    $p_{l,1}' = (1 - 0.5 \times \alpha_2) p_l$ *and* $p_{l,2}' = (1 + 0.5 \times \alpha_2) p_l$.
    *If* $(q_{l,1} - q_{l,2}) \geq \alpha_2 q_l$, *then*
    $q_{l,1}' = q_{l,1}$ *and* $q_{l,2}' = q_{l,2}$.
    *Otherwise,*
    $q_{l,1}' = (1 + 0.5 \times \alpha_2) q_l$ *and* $q_{l,2}' = (1 - 0.5 \times \alpha_2) q_l$.

After obtaining $p'_{l,1}$, $p'_{l,2}$, $q'_{l,1}$ and $q'_{l,2}$, the DCT coefficients in the corresponding fragments are modified by

$$\chi'^{f}_{l,j}(k) = \chi^{f}_{l,j}(k) \times \frac{p'_{l,j}}{p_{l,j}} \tag{16}$$

$$\chi'^{r}_{l,j}(k) = \chi^{r}_{l,j}(k) \times \frac{q'_{l,j}}{q_{l,j}} \tag{17}$$

where $j = 1, 2$. From (7), (8), (16) and (17), it is obvious that

$$E\left(\left|\chi'^{f}_{l,j}(k)\right|\right) = p'_{l,j} \tag{18}$$

$$E\left(\left|\chi'^{r}_{l,j}(k)\right|\right) = q'_{l,j}. \tag{19}$$

Similarly, we can insert the same watermark bit into all the selected DCT frame pairs from the same host audio segment. After that, the watermarked audio segment is constructed by using inverse discrete cosine transform (IDCT).

It is worthwhile to note that in the proposed embedding algorithm, the positive constant $\alpha_2$ is introduced to increase the robustness against attacks, which will become clearer in the Section II-B. However, $\alpha_2$ should be kept small to ensure high perceptual quality.

We would also like to note that the new embedding algorithm is purposely designed to ensure that the selection criterion in (14) can be applied in the decoding process to find the watermarked frame pairs from the watermarked signal. This can be explained as follows. We assume that the $l$th frame pair $\{\chi^{f}_{l}(k), \chi^{r}_{l}(k)\}$ is selected to embed a watermark bit. It results from (14) that $\tilde{p}_l \leq 0$ and $\tilde{q}_l \leq 0$. Let $\{\chi'^{f}_{l}(k), \chi'^{r}_{l}(k)\}$ be the modified (or watermarked) counterpart of $\{\chi^{f}_{l}(k), \chi^{r}_{l}(k)\}$, and denote

$$p'_l = E\left(\left|\chi'^{f}_{l}(k)\right|\right) \tag{20}$$

$$q'_l = E\left(\left|\chi'^{r}_{l}(k)\right|\right). \tag{21}$$

Similar to (11)–(13), we define

$$\tilde{p}'_l = |p'_{l,1} - p'_{l,2}| - \alpha_1 p'_l \tag{22}$$

$$\tilde{q}'_l = |q'_{l,1} - q'_{l,2}| - \alpha_1 q'_l \tag{23}$$

where

$$p'_l = \frac{p'_{l,1} + p'_{l,2}}{2} \quad \text{and} \quad q'_l = \frac{q'_{l,1} + q'_{l,2}}{2}. \tag{24}$$

As shown in the proposed embedding rule, different equations are used for watermark embedding, depending on the watermark bit to be embedded. For the fragments $\{\chi^{f}_{l,1}(k), \chi^{f}_{l,2}(k)\}$, two sets of equations are used, which are 1) $p'_{l,1} = p_{l,1}$ and $p'_{l,2} = p_{l,2}$, and 2) $p'_{l,1} = (1 \pm 0.5 \times \alpha_2)p_l$ and $p'_{l,2} = (1 \mp 0.5 \times \alpha_2)p_l$. In the first scenario, it is obvious from (11) and (24) that $p'_l = p_l$. Then, it follows from (12) and (22) that $\tilde{p}'_l = \tilde{p}_l$, i.e., $\tilde{p}_l \leq 0$ results in $\tilde{p}'_l \leq 0$. In the second scenario, it yields from (24) that

$$p'_l = \frac{(1 \pm 0.5 \times \alpha_2)p_l + (1 \mp 0.5 \times \alpha_2)p_l}{2} = p_l. \tag{25}$$

Then, from (15), (22) and (25), it follows

$$\tilde{p}'_l = |(1 \pm 0.5 \times \alpha_2)p_l - (1 \mp 0.5 \times \alpha_2)p_l| - \alpha_1 p_l$$
$$= (\alpha_2 - \alpha_1)p_l < 0.$$

Hence, the embedding rule guarantees that $\tilde{p}_l \leq 0$ leads to $\tilde{p}'_l \leq 0$. Similarly, for the fragments $\{\chi^{r}_{l,1}(k), \chi^{r}_{l,2}(k)\}$, we can show that $q'_l = q_l$, and $\tilde{q}_l \leq 0$ yields $\tilde{q}'_l \leq 0$. Since $p'_l = p_l$, $q'_l = q_l$, and $\tilde{p}_l \leq 0$ and $\tilde{q}_l \leq 0$ always result in $\tilde{p}'_l \leq 0$ and $\tilde{q}'_l \leq 0$, then the selection criterion (14) can be employed in the decoding process to identify the watermarked frame pairs.

### B. Watermark Decoding

First, we use the segmenting procedure utilized in the embedding process to partition the watermarked audio signal. Thus, for a given segment of the watermarked audio signal, we can obtain its DCT front and rear subsegments $\{\chi'^{f}(k), \chi'^{r}(k)\}$, frame pairs $\{\chi'^{f}_{i}(k), \chi'^{r}_{i}(k)\}$, and the associated fragments $\{\chi'^{f}_{i,j}(k), \chi'^{r}_{i,j}(k)\}$, where $i = 1, 2, \ldots, R$ and $j = 1, 2$. Here, the PN sequence $p(n)$ is used as a secret key in the construction of the fragments $\{\chi'^{f}_{i,j}(k), \chi'^{r}_{i,j}(k)\}$. Before attempting to extract a watermark bit from a frame pair, say the $l$th frame pair, it is essential to find whether the concerned frame pair contains a watermark or not. To do so, we first use (18)–(21) to compute $p'_{l,j}$, $q'_{l,j}$, $p'_l$ and $q'_l$, where $j = 1, 2$. Then, one can obtain $\tilde{p}'_l$ and $\tilde{q}'_l$ from (22) and (23). As we analyzed at the end of the previous subsection, the selection criterion (14) can be applied at the decoding stage to identify the watermarked frame pairs. Specifically, if

$$\tilde{p}'_l \leq 0, \quad \tilde{q}'_l \leq 0, \quad \text{and} \quad \min\{p'_l, q'_l\} \geq \theta_{th} \tag{26}$$

the $l$th frame pair $\left\{X'^{f}_{l}(k), X'^{r}_{l}(k)\right\}$ contains a watermark. Based on (26), all the frame pairs containing a watermark can be determined. Then the embedded watermark can be extracted using the watermark extraction approach to be presented next. If none of the frame pairs associated with the given segment of the watermarked audio signal satisfies (26), the concerned audio segment does not contain a watermark and watermark extraction will not be carried out.

Assume that the $l$th frame pair contains a watermark. To extract the watermark bit, we define

$$\left|\tilde{\chi}'^{r}_{l,j}(k)\right| = \left|\chi'^{r}_{l,j}(k)\right| + (p'_l - q'_l).$$

From the above equation, it follows

$$E\left(\left|\tilde{\chi}'^{r}_{l,j}(k)\right|\right) = E\left(\left|\chi'^{r}_{l,j}(k)\right|\right) + (p'_l - q'_l)$$
$$= q'_{l,j} + (p'_l - q'_l) = q'_{l,j} + (p_l - q_l) \tag{27}$$

where $j = 1, 2$. The derivation of the above last equation results from the fact that $p'_l = p_l$ and $q'_l = q_l$. Moreover, let

$$r_{l,j} = E\left(\left|\chi'^{f}_{l,j}(k)\right|\right) - E\left(\left|\tilde{\chi}'^{r}_{l,j}(k)\right|\right)$$
$$= p'_{l,j} - (q'_{l,j} + p_l - q_l) = (p'_{l,j} - p_l) + (q_l - q'_{l,j}) \tag{28}$$

where $j = 1, 2$. Now, we consider two cases.

*1) Case 1: The Embedded Watermark is "0":* As shown in the embedding rule, if $(p_{l,1} - p_{l,2}) \geq \alpha_2 p_l$, then $p'_{l,1} = p_{l,1}$, which yields

$$p'_{l,1} - p_l = p_{l,1} - p_l = p_{l,1} - \frac{p_{l,1} + p_{l,2}}{2}$$
$$= \frac{p_{l,1} - p_{l,2}}{2} \geq 0.5 \times \alpha_2 p_l.$$

TABLE I
HOST AUDIO SIGNALS USED IN SIMULATIONS

| Host signals | Genres |
|---|---|
| $S_{01} \sim S_{10}$ | Western pop music |
| $S_{11} \sim S_{20}$ | Eastern classical music |
| $S_{21} \sim S_{30}$ | South asian folk music |
| $S_{31} \sim S_{40}$ | Subcontinent country music |
| $S_{41} \sim S_{50}$ | Speeches |

On the other hand, if $(p_{l,1} - p_{l,2}) < \alpha_2 p_l$, then $p'_{l,1} = (1 + 0.5 \times \alpha_2)p_l$, which yields $p'_{l,1} - p_l = 0.5 \times \alpha_2 p_l$. Thus, $p'_{l,1} - p_l \geq 0.5 \times \alpha_2 p_l$ holds for this case. Similarly, it can be shown that $q_l - q'_{l,1} \geq 0.5 \times \alpha_2 q_l$ also holds for this case. Hence, we can obtain from (28) that

$$r_{l,1} = (p'_{l,1} - p_l) + (q_l - q'_{l,1})$$
$$\geq (0.5 \times \alpha_2 p_l) + (0.5 \times \alpha_2 q_l) > 0. \quad (29)$$

Following the same way, one can show

$$r_{l,2} = (p'_{l,2} - p_l) + (q_l - q'_{l,2})$$
$$\leq (-0.5 \times \alpha_2 p_l) + (-0.5 \times \alpha_2 q_l) < 0. \quad (30)$$

*2) Case 2: The Embedded Watermark is "1":* For this case, one can verify that

$$r_{l,1} \leq -0.5 \times \alpha_2 (p_l + q_l) < 0 \quad (31)$$
$$r_{l,2} \geq 0.5 \times \alpha_2 (p_l + q_l) > 0. \quad (32)$$

Based on (29)–(32), if $r_{l,1} > 0$ and $r_{l,2} < 0$, then the watermark bit "0" is extracted from the $l$th frame pair. Otherwise, the extracted watermark bit is "1." After extracting watermark bits from all frame pairs satisfying (26), the watermark bit embedded in the associated audio segment is determined by majority rule. By this way, all embedded watermarks can be extracted from the watermarked audio signal.

From the principle of watermark extraction, it is clear that in the absence of attacks, watermarks can be correctly extracted under any small value of $\alpha_2$ so long as $\alpha_2 > 0$. In the presence of attacks, increasing $\alpha_2$ can improve robustness, at the price of lowering perceptual quality. Furthermore, embedding the same watermark bit multiple times in one audio segment makes the statistical decision making more effective and provides error correcting functionality [2]. As a result, it helps to further enhance the robustness against attacks.

## III. SIMULATION RESULTS

In this section, simulation examples are provided to demonstrate the performance of the proposed scheme. In the simulations, we use 50 randomly selected mono-channel audio clips belonging to five different groups as host signals (see Table I). All these audio clips have a duration of 10 seconds. They are sampled at the rate of 44.1 kHz, quantized with 16 bits, and then segmented.

*1) Example 1:* A practically feasible watermarking method should be robust to conventional attacks while maintaining high perceptual quality. The imperceptibility of the proposed watermarking scheme depends on the watermarking parameters $\alpha_1$ and $\alpha_2$. To asses the imperceptibility of our scheme, we employ
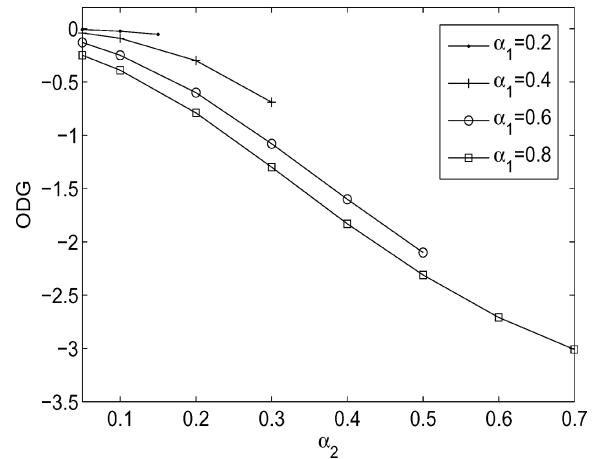


Fig. 2. ODG versus $\alpha_2$ with $\alpha_1 = 0.2, 0.4, 0.6$, and $0.8$, respectively.

the Perceptual Evaluation of Audio Quality (PEAQ) algorithm [25], as used in [3], [26] and [27]. The PEAQ algorithm compares the quality of the host (unwatermarked) signal with its watermarked counterpart and returns a parameter called objective difference grade (ODG) ranging between $-4$ and $0$. The perceptual quality improves with the increase of the ODG value.

We use the above audio clips to calculate the ODG values of the proposed watermarking scheme under different $\alpha_1$ and $\alpha_2$. The audio clips are segmented in such a way that each segment contains 8820 samples. Also, we choose $M = 15$, $R = 40$, and $f_{\text{limit}} = 6$ kHz. As shown in Fig. 2, given an $\alpha_1$, the ODG value decreases with the increase of $\alpha_2$. Similarly, for a fixed $\alpha_2$, the ODG value also decreases with the rise of $\alpha_1$. This figure is helpful for the selection of $\alpha_1$ and $\alpha_2$. For example, the ODG value of $-0.3$ ensuring high imperceptibility [3], [27] can be obtained by choosing $\alpha_1 = 0.4$ and $\alpha_2 = 0.2$. In addition, the following simulation examples will show that this set of $\alpha_1$ and $\alpha_2$ are also suitable for our scheme to have high robustness against attacks.

*2) Example 2:* This example evaluates the robustness of the proposed watermarking scheme against common attacks. The robustness is measured by bit error rate (BER), which is defined as follows:

$$\text{BER} = \frac{\text{Number of watermarks incorrectly extracted}}{\text{Number of watermarks embedded}} \times 100\%.$$

The following common attacks are used in the evaluation of robustness:

*Closed-loop attack*: The watermarks are extracted from the watermarked signals without any attacks.

*Re-quantization attack*: Each sample of the watermarked signals is re-quantized from 16 bits to 8 bits.

*Noise attack*: Random noise is added to the watermarked signals, where the ratio of the watermarked signal to noise is 20 dB.

*Amplitude attack*: The amplitudes of the watermarked signals are enlarged by 1.8 times.

*MP3 attack*: MPEG 1 Layer III compression is performed on the watermarked signals.

*AAC attack*: MPEG 4 advanced audio coding based compression is performed on the watermarked signals.

TABLE II
BERs OF THE PROPOSED SCHEME AND THE METHOD IN [3], WHERE
EMBEDDING RATE IS 5 bps AND $\mathrm{ODG} = -0.3$ FOR BOTH METHODS

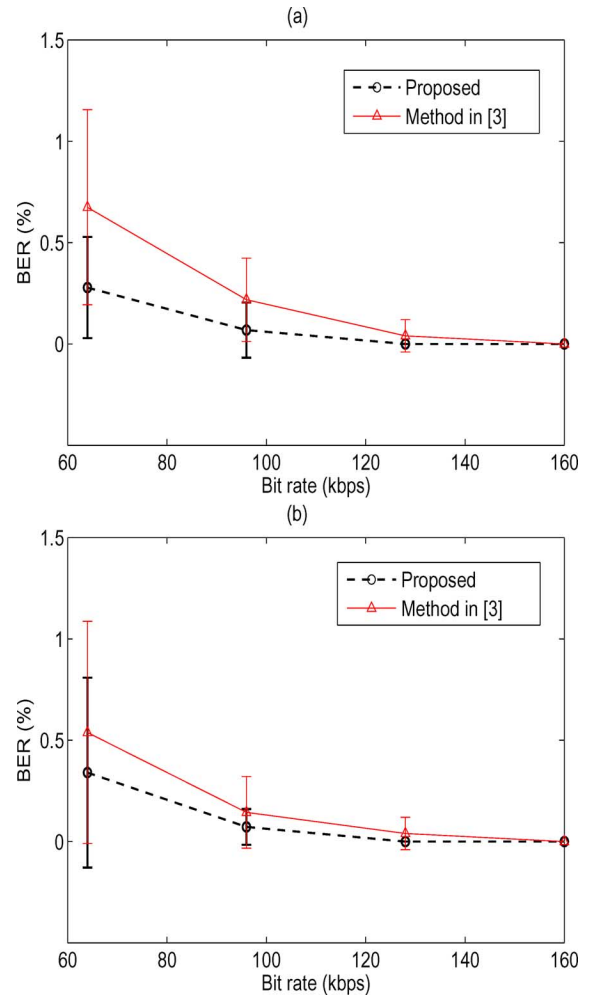| Attacks | Host signals | BER (%) | |
|---|---|---|---|
| | | Method in [3] | Proposed method |
| Closed-loop | $S_{01} \sim S_{10}$ | 0 | 0 |
| | $S_{11} \sim S_{20}$ | 0 | 0 |
| | $S_{21} \sim S_{30}$ | 0 | 0 |
| | $S_{31} \sim S_{40}$ | 0 | 0 |
| | $S_{41} \sim S_{50}$ | 0 | 0 |
| Re-quantization | $S_{01} \sim S_{10}$ | 1.7 | 0.2 |
| | $S_{11} \sim S_{20}$ | 0.6 | 0 |
| | $S_{21} \sim S_{30}$ | 1.8 | 0 |
| | $S_{31} \sim S_{40}$ | 0.8 | 0 |
| | $S_{41} \sim S_{50}$ | 6.4 | 0 |
| Noise | $S_{01} \sim S_{10}$ | 3.0 | 0.6 |
| | $S_{11} \sim S_{20}$ | 2.3 | 0.4 |
| | $S_{21} \sim S_{30}$ | 1.5 | 0 |
| | $S_{31} \sim S_{40}$ | 0.6 | 0 |
| | $S_{41} \sim S_{50}$ | 5.4 | 0.4 |
| Amplitude | $S_{01} \sim S_{10}$ | 0 | 0 |
| | $S_{11} \sim S_{20}$ | 0 | 0 |
| | $S_{21} \sim S_{30}$ | 0 | 0 |
| | $S_{31} \sim S_{40}$ | 0 | 0 |
| | $S_{41} \sim S_{50}$ | 0 | 0 |
| MP3 (128kbps) | $S_{01} \sim S_{10}$ | 0 | 0 |
| | $S_{11} \sim S_{20}$ | 0.2 | 0 |
| | $S_{21} \sim S_{30}$ | 0 | 0 |
| | $S_{31} \sim S_{40}$ | 0 | 0 |
| | $S_{41} \sim S_{50}$ | 0 | 0 |
| AAC (128kbps) | $S_{01} \sim S_{10}$ | 0 | 0 |
| | $S_{11} \sim S_{20}$ | 0.2 | 0 |
| | $S_{21} \sim S_{30}$ | 0 | 0 |
| | $S_{31} \sim S_{40}$ | 0 | 0 |
| | $S_{41} \sim S_{50}$ | 0 | 0 |
| Re-sampling | $S_{01} \sim S_{10}$ | 0.8 | 0.2 |
| | $S_{11} \sim S_{20}$ | 12.0 | 1.6 |
| | $S_{21} \sim S_{30}$ | 1.1 | 0.2 |
| | $S_{31} \sim S_{40}$ | 0.5 | 0 |
| | $S_{41} \sim S_{50}$ | 9.3 | 0.6 |
| HPF (100Hz) | $S_{01} \sim S_{10}$ | 0.2 | 0 |
| | $S_{11} \sim S_{20}$ | 1.0 | 0 |
| | $S_{21} \sim S_{30}$ | 0 | 0 |
| | $S_{31} \sim S_{40}$ | 0.4 | 0 |
| | $S_{41} \sim S_{50}$ | 1.9 | 0 |
| LPF (8kHz) | $S_{01} \sim S_{10}$ | 0.4 | 0 |
| | $S_{11} \sim S_{20}$ | 0.6 | 0 |
| | $S_{21} \sim S_{30}$ | 0.6 | 0 |
| | $S_{31} \sim S_{40}$ | 0 | 0 |
| | $S_{41} \sim S_{50}$ | 1.0 | 0 |



Fig. 3. BERs under MP3 and AAC attacks versus compression bit rate, together with the corresponding standard deviations among the tested genres. (a) MP3 attack. (b) AAC attack.

*Re-sampling attack*: The watermarked signals are down-sampled to 16 kHz and then up-sampled back to 44.1 kHz (i.e., $44.1 \text{ kHz} \rightarrow 16 \text{ kHz} \rightarrow 44.1 \text{ kHz}$).
*High-pass filtering (HPF)*: High-pass filter with 100 Hz cutoff frequency is applied to the watermarked signals.
*Low-pass filtering (LPF)*: Low-pass filter with 8 kHz cutoff frequency is applied to the watermarked signals.

Since the proposed embedding and decoding scheme is built upon the patchwork concept, we compare it with the latest patchwork method in [3]. Precisely, we compare the robustness of both approaches under the same embedding rate of 5 bps and the same perceptual quality with $\mathrm{ODG} = -0.3$. Since the embedding rate of the method in [3] is signal dependent, we adjust the segment length for each audio clip such that the average embedding rate over all audio clips is 5 bps. For the proposed scheme, the number of samples in a segment is fixed at 8820 for all audio clips, which results in 5 bps embedding

rate. We choose $\mathrm{ODG} = -0.3$ as it guarantees high perceptual quality for both approaches. Other simulation parameters for the proposed scheme are $M = 15$, $R = 40$, $\alpha_1 = 0.4$, $\alpha_2 = 0.2$, $\theta_{th} = 0.0005$, and $f_{\text{limit}} = 6 \text{ kHz}$. For the method in [3], other simulation parameters are $\gamma_{min} = 1.03$, $\mathrm{CGF} = 0.8$, $\alpha = 0.05$ and $Q = 0.03$, which are the same as those used in [3].

Table II shows the BERs of the new scheme and the method in [3], under the aforementioned common attacks. Here, the bit rate of 128 kbps is used for MP3 and AAC attacks. It can be seen that both approaches achieves 0% bit error rate under closed-loop attack and amplitude attack. However, our watermarking scheme outperforms the one in [3] under all other attacks. In particular, the proposed scheme also yields 0% bit error rate under MP3, AAC, HPF, and LPF attacks.

*3) Example 3:* In the third example, we further evaluate the robustness of our scheme and the method in [3] against MP3 and AAC attacks at different compression bit rates: 64 kbps, 96 kbps, 128 kbps, and 160 kbps. These compression bit rates are commonly used in practical applications. Other simulation parameters for both methods are the same as those in Example 2. Fig. 3 shows the BERs versus compression bit rate, as well as the corresponding standard deviations among the tested genres.

One can see from Fig. 3 that as anticipated, BERs decrease with the rise of the compression bit rate for both approaches. However, our scheme performs better than the other method under both compression attacks and at different compression bit rates. Furthermore, it is shown that there are BER deviations among different genres for both approaches and the deviations tend to decrease when the compression bit rate increases. However, our method results in smaller deviations than the method in [3] does at all considered compression bit rates.

It should be noted that the method in [3] needs explicit knowledge about which segments of the watermarked audio signal contain watermarks at the decoding end. Without this information, watermarks cannot be correctly extracted from the watermarked signal by this method. In the simulations, we assume that this information is known at the decoding end for the method in [3]. However, to our best knowledge, identifying the watermarked segments encountered in [3] is still an open problem. Therefore, the practical usage of this method is limited. On the contrary, our scheme does not need any information to identify the watermarked frame pairs in the decoding process.

## IV. Conclusion

In this paper, we propose a robust patchwork-based watermarking scheme for audio signals, which inserts watermarks into audio signals by modifying their DCT coefficients. In the proposed scheme, watermarks are only embedded into suitable DCT frame pairs to ensure high imperceptibility. Besides, the embedding algorithm is designed in such a way that the criterion used in the embedding process to select suitable DCT frame pairs can be applied at the decoding stage to find the watermarked DCT frame pairs. Thus, unlike the latest patchwork method in [3], our scheme does not require any additional information to discover the watermarked DCT frame pairs. Furthermore, high robustness is ensured by the mechanism of the proposed scheme together with the usage of selected frequency region and multiple watermark embedding. The new scheme is also secure due to the necessity of using a secret key in the decoding process. The superior performance of our scheme is demonstrated by simulation examples, in comparison with the method in [3].
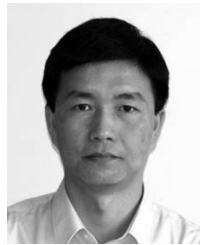
## References

[1] X. Wang, W. Qi, and P. Niu, "A new adaptive digital audio watermarking based on support vector regression," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 15, no. 8, pp. 2270–2277, Nov. 2007.

[2] I. K. Yeo and H. J. Kim, "Modified patchwork algorithm: A novel audio watermarking scheme," *IEEE Trans. Speech Audio Process.*, vol. 11, no. 4, pp. 381–386, Jul. 2003.

[3] N. K. Kalantari, M. A. Akhaee, S. M. Ahadi, and H. Amindavar, "Robust multiplicative patchwork method for audio watermarking," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 17, no. 6, pp. 1133–1141, Aug. 2009.

[4] W.-N. Lie and L.-C. Chang, "Robust and high-quality time-domain audio watermarking based on low-frequency amplitude modification," *IEEE Trans. Multimedia*, vol. 8, no. 1, pp. 46–59, Feb. 2006.

[5] H. Wang, R. Nishimura, Y. Suzuki, and L. Miao, "Fuzzy self-adaptive digital audio watermarking based on time-spread echo hiding," *Appl. Acoust.*, vol. 69, no. 10, pp. 868–874, Oct. 2008.

[6] O. T.-C. Chen and W.-C. Wu, "Highly robust, secure, and perceptual-quality echo hiding scheme," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 16, no. 3, pp. 629–638, Mar. 2008.

[7] L. Luo, Z. Chen, M. Chen, X. Zeng, and Z. Xiong, "Reversible image watermarking using interpolation technique," *IEEE Trans. Inf. Forensics Security*, vol. 5, no. 1, pp. 187–193, Mar. 2010.

[8] C. H. Huang, S. C. Chuang, Y. L. Huang, and J. L. Wu, "Unseen visible watermarking: A novel methodology for auxiliary information delivery via visual contents," *IEEE Trans. Inf. Forensics Security*, vol. 4, no. 2, pp. 193–206, Jun. 2009.

[9] H. Huang, C. Yang, and W. Hsu, "A video watermarking technique based on pseudo-3-D DCT and quantization index modulation," *IEEE Trans. Inf. Forensics Security*, vol. 5, no. 4, pp. 625–637, Dec. 2010.

[10] M. Arnold, "Audio watermarking: Features, applications and algorithm," in *Proc. IEEE Int. Conf. Multimedia Expo*, 2000, vol. 2, pp. 1013–1016.

[11] Z. Liu and A. Inoue, "Audio watermarking techniques using sinusoidal pattern based on pseudorandom sequence," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 8, pp. 801–812, Aug. 2003.

[12] A. Valizadeh and Z. J. Wang, "Correlation-and-bit-aware spread spectrum embedding for data hiding," *IEEE Trans. Inf. Forensics Security*, vol. 6, no. 2, pp. 267–282, Jun. 2011.

[13] N. Cvejic and T. Seppänen, "Spread spectrum audio watermarking using frequency hopping and attack characterization," *Signal Process.*, vol. 84, pp. 207–213, 2004.

[14] B.-S. Ko, R. Nishimura, and Y. Suzuki, "Time-spread echo method for digital audio watermarking," *IEEE Trans. Multimedia*, vol. 7, no. 2, pp. 212–221, Apr. 2005.

[15] Y. Xiang, D. Peng, I. Natgunanathan, and W. Zhou, "Effective pseudonoise sequence and decoding function for imperceptibility and robustness enhancement in time-spread echo based audio watermarking," *IEEE Trans. Multimedia*, vol. 13, no. 1, pp. 2–13, Feb. 2011.

[16] Y. Xiang, I. Natgunanathan, D. Peng, W. Zhou, and S. Yu, "A dual-channel time-spread echo method for audio watermarking," *IEEE Trans. Inf. Forensics Security*, vol. 7, no. 2, pp. 383–392, Apr. 2012.

[17] Y. Erfani and S. Siahpoush, "Robust audio watermarking using improved TS echo hiding," *Digital Signal Process.*, vol. 19, pp. 809–814, 2009.

[18] H. J. Kim and Y. H. Choi, "A novel echo-hiding scheme with backward and forward kernels," *IEEE Trans. Circuit Syst. Video Technol.*, vol. 13, no. 8, pp. 885–889, Aug. 2003.

[19] S. Kirbiz and B. Gunsel, "Robust audio watermark decoding by supervised learning," in *Proc. IEEE Int. Conf. Acoust,, Speech, Signal Process.*, 2006, pp. 761–764.

[20] D. Lakshmi, R. Ganesh, R. Marni, R. Prakash, and P. Arulmozhivarman, "SVM based effective watermarking scheme for embedding binary logo and audio signals in images," in *Proc. IEEE Region 10 Conf.*, 2008, pp. 1–5.

[21] H. Kang, K. Yamaguchi, B. Kurkoski, K. Yamaguchi, and K. Kobayashi, "Full-index-embedding patchwork algorithm for audio watermarking," *IEICE Trans. Inf. Syst.*, vol. E91-D, no. 11, pp. 2731–2734, Nov. 2008.

[22] W. Bender, D. Gruhl, N. Morimoto, and A. Lu, "Techniques for data hiding," *IBM Syst. J.*, vol. 35, no. 3&4, pp. 313–335, 1996.

[23] K. Solanki, N. Jacobsen, U. Madhow, B. S. Manjunath, and S. Chandrasekaran, "Robust image-adaptive data hiding based on erasure and error correction," *IEEE Trans. Image Process.*, vol. 13, no. 12, pp. 1627–1639, Dec. 2004.

[24] J. L. Wu and J. Shin, "Discrete cosine transform in error control coding," *IEEE Trans. Commun.*, vol. 43, no. 5, pp. 1857–1861, May 1995.

[25] "Rec.B. S. 1387: Methods for objective measurements of perceived audio quality," Int. Telecomm. Union, Geneva, Switzerland, Rec. B.S. 1387, 2001.

[26] C. Baras, N. Moreau, and P. Dymarski, "Controlling the inaudibility and maximizing the robustness in an audio annotation watermarking system," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 14, no. 5, pp. 1772–1782, Sep. 2006.

[27] N. T. H. Lien, H. Nobuhara, F. Dong, and K. Hirota, "Two channel digital watermarking for music based on exponential time-spread echo kernel," *Signal, Image, Video Process.*, vol. 3, no. 2, pp. 115–126, Feb. 2009.

**Iynkaran Natgunanathan** received the B.Sc. Eng. (Honors) degree in electronics and telecommunication engineering from the University of Moratuwa, Katubedda, Moratuwa, Sri Lanka, in 2007. He is currently pursuing the Ph.D. degree with the School of Engineering, Deakin University, Geelong, Australia.

From 2006 to 2008, he was a Software Engineer at Millennium Information Technology (Pvt) Ltd., Malabe, Sri Lanka. His research interests include digital watermarking, psychoacoustics, audio/speech/image processing, telecommunication, and robotics.

**Yong Xiang** received the B.E. and M.E. degrees from the University of Electronic Science and Technology of China, Chengdu, in 1983 and 1989, respectively, and the Ph.D. degree from The University of Melbourne, Melbourne, Australia, in 2003.

He was with the Southwest Institute of Electronic Equipment of China, Chengdu, from 1983 to 1986. In 1989, he joined the University of Electronic Science and Technology of China, where he was a Lecturer from 1989 to 1992 and an Associate Professor from 1992 to 1997. He was a Senior Communications Engineer with Bandspeed, Inc., Melbourne, Australia, from 2000 to 2002. He is currently an Associate Professor with the School of Information Technology at Deakin University, Melbourne, Australia. His research interests include blind signal/system estimation, communication signal processing, biomedical signal processing, information and network security, speech and image processing, and pattern recognition.

**Yue Rong** (S'03–M'06–SM'11) received the B.E. degree from Shanghai Jiao Tong University, Shanghai, China, the M.Sc. degree from the University of Duisburg-Essen, Duisburg, Germany, and the Ph.D. degree (*summa cum laude*) from Darmstadt University of Technology, Darmstadt, Germany, all in electrical engineering, in 1999, 2002, and 2005, respectively.

From April 2001 to October 2001, he was a Research Assistant at the Fraunhofer Institute of Microelectronic Circuits and Systems, Duisburg. From October 2001 to March 2002, he was with Nokia, Ltd., Bochum, Germany. From November 2002 to March 2005, he was a Research Associate at the Department of Communication Systems in the University of Duisburg-Essen. From April 2005 to January 2006, he was with the Institute of Telecommunications at Darmstadt University of Technology, as a Research Associate. From February 2006 to November 2007, he was a Postdoctoral Researcher with the Department of Electrical Engineering, University of California, Riverside. Since December 2007, he has been with the Department of Electrical and Computer Engineering, Curtin University of Technology, Perth, Australia, where he is now a Senior Lecturer. His research interests include signal processing for communications, wireless communications, wireless networks, applications of linear algebra and optimization methods, and statistical and array signal processing. He has coauthored more than 70 referred IEEE journal and conference papers.

Dr. Rong received the Best Paper Award at the Third International Conference on Wireless Communications and Signal Processing, Nanjing, China, in 2011, the Best Paper Award at the 16th Asia-Pacific Conference on Communications, Auckland, New Zealand, in 2010, the 2010 Young Researcher of the Year Award of the Faculty of Science and Engineering at Curtin University, the 2004 Chinese Government Award for Outstanding Self-Financed Students Abroad (China), and the 2001–2002 DAAD/ABB Graduate Sponsoring Asia Fellowship (Germany). He is an Editor of the IEEE WIRELESS COMMUNICATIONS LETTERS, a Guest Editor of the IEEE JOURNAL OF SELECTED AREAS IN COMMUNICATIONS special issue on Theories and Methods for Advanced Wireless Relays to be published in 2012, a Guest Editor of the *EURASIP JASP* Special Issue on Signal Processing Methods for Diversity and Its Applications to be published in 2012, and has served as a TPC member for IEEE ICC, IEEE ICCC, WCSP, IWMC, and ChinaCom.

**Wanlei Zhou** (M'02–SM'09) received the B.Eng. and M.Eng. degrees from the Harbin Institute of Technology, Harbin, China, in 1982 and 1984, respectively, the Ph.D. degree from The Australian National University, Canberra, Australia, in 1991, and the D.Sc. degree from Deakin University, Victoria, Australia, in 2002.

He is currently the Chair Professor of Information Technology and the Head of School of Information Technology, Faculty of Science and Techn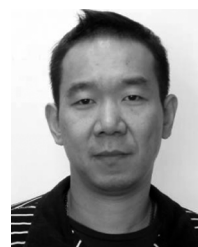ology, Deakin University, Australia. His research interests include distributed and parallel systems, network and system security, mobile computing, bioinformatics, and e-learning. He has published more than 200 papers in refereed international journals and refereed international conferences proceedings.

Prof. Zhou has served as General Chair, Steering Committee Chair, PC Chair, Publication Chair and Session Chair, for more than 20 international conferences.

**Song Guo** (M'05–SM'11) received the Ph.D. degree in computer science from the University of Ottawa, Ottawa, ON, Canada, in 2005.

He is currently a Senior Associate Professor at School of Computer Science and Engineering, University of Aizu, Aizuwakamatsu, Japan. His research interests are mainly in the areas of protocol design and performance analysis for computer and telecommunication networks, currently focusing on network modeling, cross-layer design and optimization, security analysis, and performance evaluation of wireless, mobile, and personal networks.

Dr. Guo is an Associate Editor of the IEEE TRANSACTIONS ON PARALLEL AND DISTRIBUTED SYSTEMS and an Associate Editor of *Wireless Communications and Mobile Computing*.