Reversible Watermarking on Stereo Audio Signals by Exploring Inter-Channel Correlation

Yuanxin Wu, CAS Key Laboratory of Electro-magnetic Space Information, Chinese Academy of Sciences, University of Science and Technology of China, Hefei, China
Wen Diao, CAS Key Laboratory of Electro-magnetic Space Information, Chinese Academy of Sciences, University of Science and Technology of China, Hefei, China
Dongdong Hou, CAS Key Laboratory of Electro-magnetic Space Information, Chinese Academy of Sciences, University of Science and Technology of China, Hefei, China
Weiming Zhang*, CAS Key Laboratory of Electro-magnetic Space Information, Chinese Academy of Sciences, University of Science and Technology of China, Hefei, China

ABSTRACT

A new reversible watermarking algorithm on stereo audio signals is proposed in this paper. By utilizing correlations between two channels of audio signal, we segment one channel based on another one according to the smoothness. For each segmented sub-host sequence, we estimate its capacity and the corresponding embedding distortion firstly, and then select the optimal combinations of sub-host sequences for embedding. Experimental results indicate that the proposed algorithm can improve SNR (signal to noise ratio) for various kinds of capacity.

KEYWORDS

Audio, Reversible watermarking, Stereo audio signals

1. INTRODUCTION

With the rapid development of multimedia and network, the amount of information storage becomes larger and larger. At the same time, editing and copying is so convenient, which speeds up the spread of information. Currently, a lot of digital works are suffering from illegal acquiring and malicious tampering, among which audios are the popular ones. Integrity protection and ownership rights certification of audio files have attracted great attentions, which can be realized with watermarking.

There are two types of watermarking, robust watermarking and fragile watermarking to protect audio files. Robust watermarking [1]-[4] is used to label copyright information so that protecting copyright. On the contrary, fragile watermarking [5]-[7] is usually used for content integrity authentication, which is asked to be sensitive to the slight change so that editing the content slightly will be detected. Reversible watermarking mainly is used for fragile watermarking, which can restore

* Corresponding author.
E-mail address: zhangwm@ustc.edu.cn (W.Zhang)

both the embedded watermark and the host signal. The reversibility is very important in some special situations, such as high quality music, legal evidence, military intelligence and criminal investigation.

At the early age, Barton [8] proposed the idea of reversible data hiding (RDH). Later on, many efficient methods for images spring up. The current algorithms are divided into five mainstream techniques: lossless compression based schemes [9]-[11], expansion based schemes [12]-[18], content adaptive schemes [19]-[20] and integer transform schemes [21]-[23].

Following the development of RDH in the image field, a lot of efficient reversible watermarking algorithms have been proposed. In the early days, Michiel et al. [24] proposed the reversible audio watermarking, which uses the redundant bits of audio coding to encode the watermark information, and then recover the host signal by restoring the original dynamic range in the decoder. Later on, many methods are proposed in time domain [25] [26] [27] [28] [29], compressed domain [30]-[31] and frequency domain [32] [33]. Yan et al. [25] refer to the method of extending the prediction error in the image reversible hiding proposed by Tian [12] to construct the appropriate prediction model to realize the reversible data hiding. Bradley et al. [27] proposed a high capacity reversible audio watermarking method based on the generalized reversible integer transform. Xiang et al. [29] proposed an alterable prediction order data hiding method based on non-causal prediction, in which the double-embedding strategy in image data hiding is used to divide audio signal for two sets. In the compressed domain, Li et al.[30] designed an entropy coding algorithm, the perceptually unimportant indices in one segment of compressed speech bitstream are coded by the algorithm. Huang et al. [32] achieved the adaptive embedding of watermark information by processing the DCT coefficients. Particularly, using human auditory perception characteristics can achieve a better effect for reversible watermarking in the audio. Masashi Unoki et al. [34] proposed a kind of non-perceptual audio reversible watermarking based on the delay characteristics of the human ear.

All the above RDH methods are designed for single channel audio. However, to balance the hearing effect, most of the audio files we can see on the Internet are stereo audio. In this paper we proposed a RDH method for stereo audio by exploiting the correlation between the two channels.

The rest of the paper is organized as follows. Section 2 and Section 3 give the detailed description of the proposed method, including embedding and extracting procedure. The experiment results and comparisons are presented in Section 4.The paper is concluded with a discussion in Section 5.

2. BASIC METHOD ON SINGLE CHANNEL

In this paper, we use the quantized audio samples as covers, and there are two kinds of common quantization bits in standard stereo audio, 8-bit quantization and 16-bit quantization.

For ease of understanding, we give a basic method of watermarking for the right channel signal of stereo audio M firstly.

2.1 Prediction model

As shown in *Fig.1*, the length of the signal is N and x_i^R is an integer.

x_1^R	x_2^R	x_3^R	…	x_{i-1}^R	x_i^R	x_{i+1}^R	…	x_{N-1}^R	x_N^R

Fig.1 The right channel of audio M

In the right channel, all samples are divided into even set and odd set to avoid that the modified samples affect the prediction of the current sample, and two-round embedding mechanism will be adopted.

In the first round, we only embed data into the PEs of samples in the even set. To generate PEs, the present sample x_{2i}^R is predicted as:

$$\tilde{x}_{2i}^R = u_{-3}^e x_{2i-3}^R + u_{-1}^e x_{2i-1}^R + u_1^e x_{2i+1}^R + u_3^e x_{2i+3}^R \quad (1)$$

$\mathbf{u_q^e}(q = -3, -1, 1, 3)$ is calculated by solving the following linear regression problem:

$$\mathbf{X_q^e} * \mathbf{u_q^e} = \mathbf{y_q^e} \quad (2)$$

If N is even,

$$\mathbf{X_q^e} = \begin{bmatrix} x_1^R & x_3^R & x_5^R & x_7^R \\ x_3^R & x_5^R & x_7^R & x_9^R \\ \vdots & \vdots & \vdots & \vdots \\ x_{2i-3}^R & x_{2i-1}^R & x_{2i+1}^R & x_{2i+3}^R \\ \vdots & \vdots & \vdots & \vdots \\ x_{N-7}^R & x_{N-5}^R & x_{N-3}^R & x_{N-1}^R \end{bmatrix} \quad (3)$$

$$\mathbf{y_q^e} = [x_4^R \quad x_6^R \quad \cdots \quad x_{2i}^R \quad \cdots \quad x_{N-4}^R] \quad (4)$$

If N is odd,

$$\mathbf{X_q^e} = \begin{bmatrix} x_1^R & x_3^R & x_5^R & x_7^R \\ x_3^R & x_5^R & x_7^R & x_9^R \\ \vdots & \vdots & \vdots & \vdots \\ x_{2i-3}^R & x_{2i-1}^R & x_{2i+1}^R & x_{2i+3}^R \\ \vdots & \vdots & \vdots & \vdots \\ x_{N-6}^R & x_{N-4}^R & x_{N-2}^R & x_N^R \end{bmatrix} \quad (5)$$

$$\mathbf{y_q^e} = [x_4^R \quad x_6^R \quad \cdots \quad x_{2i}^R \quad \cdots \quad x_{N-3}^R] \quad (6)$$

With the least squares method, we can get

$$\mathbf{u_q^e} = (\mathbf{X_q^e}'\mathbf{X_q^e} + \mathbf{w})^{-1}\mathbf{X_q^e}'\mathbf{y_q^e} \quad (7)$$

where $\mathbf{w}$ is a regular item to avoid NAN (not a number) problem. The value of $\mathbf{w}$ is also a factor that affects the accuracy of prediction model. According to lots of experiments, we can define an empirical value for $\mathbf{w}$:

$$\begin{cases} \mathbf{w} = diag(\mathbf{I}) \\ \mathbf{I} = [1e-5 \;\; 1e-5 \;\; 1e-5 \;\; 1e-5 \;\; 1e-5] \end{cases} \quad (8)$$

In the second round, we embed in the PEs of the odd set. Note that, in this round, the samples in the even set has been modified. The present sample x_{2i+1}^R is predicted as

$$\tilde{x}_{2i+1}^R = u_{-3}^o \hat{x}_{2i-2}^R + u_{-1}^o \hat{x}_{2i}^R + u_1^o \hat{x}_{2i+2}^R + u_3^o \hat{x}_{2i+4}^R \quad (9)$$

where $\hat{x}_{2i-2}^R$, $\hat{x}_{2i}^R$, $\hat{x}_{2i+2}^R$, $\hat{x}_{2i+4}^R$ are the modified samples. The coefficients $\mathbf{u_q^o}(q = -3,-1,1,3)$ is calculated by solving the following linear regression problem

$$\mathbf{X_q^o} * \mathbf{u_q^o} = \mathbf{y_q^o} \quad (10)$$

If N is even,

$$\mathbf{X_q^o} = \begin{bmatrix} \hat{x}_2^R & \hat{x}_4^R & \hat{x}_6^R & \hat{x}_8^R \\ \hat{x}_4^R & \hat{x}_6^R & \hat{x}_8^R & \hat{x}_{10}^R \\ \vdots & \vdots & \vdots & \vdots \\ \hat{x}_{2i-2}^R & \hat{x}_{2i}^R & \hat{x}_{2i+2}^R & \hat{x}_{2i+4}^R \\ \vdots & \vdots & \vdots & \vdots \\ \hat{x}_{N-6}^R & \hat{x}_{N-4}^R & \hat{x}_{N-2}^R & \hat{x}_N^R \end{bmatrix} \quad (11)$$

$$\mathbf{y_q^o} = [x_5^R \quad x_7^R \quad \cdots \quad x_{2i+1}^R \quad \cdots \quad x_{N-3}^R] \quad (12)$$

If N is odd,

$$\mathbf{X_q^o} = \begin{bmatrix} \hat{x}_2 & \hat{x}_4 & \hat{x}_6 & \hat{x}_8 \\ \hat{x}_4 & \hat{x}_6 & \hat{x}_8 & \hat{x}_{10} \\ \vdots & \vdots & \vdots & \vdots \\ \hat{x}_{2i-2} & \hat{x}_{2i} & \hat{x}_{2i+2} & \hat{x}_{2i+4} \\ \vdots & \vdots & \vdots & \vdots \\ \hat{x}_{N-7} & \hat{x}_{N-5} & \hat{x}_{N-3} & \hat{x}_{N-1} \end{bmatrix} \quad (13)$$

$$\mathbf{y_q^o} = [x_5^R \quad x_7^R \quad \cdots \quad x_{2i+1}^R \quad \cdots \quad x_{N-4}^R] \quad (14)$$

With the least squares method, we can get

$$\mathbf{u_q^o} = (\mathbf{X_q^o}'\mathbf{X_q^o} + \mathbf{w})^{-1}\mathbf{X_q^o}'\mathbf{y_q^o} \quad (15)$$

2.2 Embedding procedure

Through the prediction model, we can get the prediction error calculated as

$$e_i^R = \tilde{x}_i^R - x_i^R \quad (16)$$

According to e_i^R, we embed the watermark bit b into the right channel of the stereo audio M as follows :

$$\hat{e}_i^R = \begin{cases} 2e_i^R + b\,, & if\ e_i^R \in [-t,t] \\ e_i^R + t + 1\,, & if\ e_i^R \in (-t,+\infty) \\ e_i^R - t, & if\ e_i^R \in (-\infty,-t) \end{cases} \quad (17)$$

$b \in \{0,1\}$ represents the watermark bit, t is a threshold deciding the capacity.

Adding the modified error $\hat{e}_i^R$ to the current sample, we can get the marked signal:

$$\hat{x}_i^R = x_i^R + \hat{e}_i^R \quad (18)$$

2.3 Extraction and restoration procedure

We can extract the watermark bit b as

$$b = \hat{e}_i^R mod2 \quad \hat{e}_i^R \in [-2t, 2t+1] \quad (19)$$

To restore the cover signal, it is necessary to recover the original prediction error firstly as:

$$e_i^R = \begin{cases} \lfloor \hat{e}_i^R/2 \rfloor, & if\ \hat{e}_i^R \in [-2t, 2t) \\ \hat{e}_i^R - t - 1, & if\ \hat{e}_i^R \in (2t, +\infty) \\ \hat{e}_i^R + t, & if\ \hat{e}_i^R \in (-\infty, -2t) \end{cases} \quad (20)$$

and we can restore the original sample as:

$$x_i^R = \hat{x}_i^R + e_i^R \quad (21)$$

3. IMPROVED METHOD WITH INTER-CHANNEL CORRELATION

We then propose an improved method that applies to stereo audio referring to the basic method.

Previous work [25]-[28] show that efficiently exploiting correlation can increase message embedding capacity in the area of reversible watermarking. There is a strong correlation between two channels in most stereo audio files. In this section, we propose a RDH method on stereo audio by using such correlation. The overview of the proposed method is shown in *Fig.2*.

We will embed data into the right channel. First, we locate smooth regions of the right channel with inter-channel prediction. And then, in such smooth regions, we generate prediction error (PE) by intra-channel prediction. Finally, watermark is reversibly embedded by modifying the histogram of the PEs.

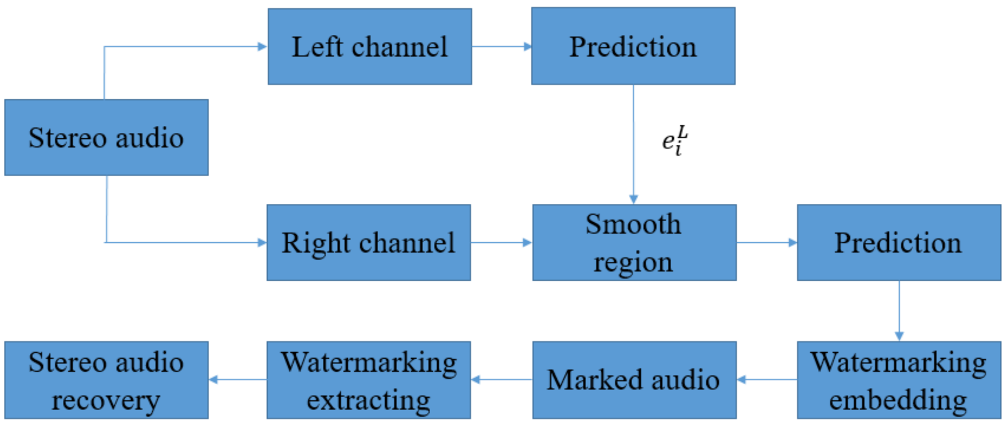

Fig.2 Framework of the reversible watermarking scheme

3.1 Correlation between two channels

We first analyze the correlation of inter channel in stereo audio. We calculate the correlation coefficient of ten stereo audio clips selected in database [37] randomly with Eq. (22).

$$q = \frac{\sum \mathbf{XY} - \frac{\sum \mathbf{X} \sum \mathbf{Y}}{N}}{\sqrt{(\sum \mathbf{X}^2 - \frac{(\sum \mathbf{X})^2}{N})(\sum \mathbf{Y}^2 - \frac{(\sum \mathbf{Y})^2}{N})}} \quad (22)$$

where $\mathbf{X}$ is the left channel signal and $\mathbf{Y}$ is the right channel. N is the length of the stereo audio. The correlation coefficients are listed in *Table 1*, which shows strong correlation between the two channels in most audio clips.

3.2 Prediction model in left channel

In *Fig.3*, the length of the left channel signal of the stereo audio M is N and x_i^L is an integer.

x_1^L	x_2^L	x_3^L	...	x_{i-1}^L	x_i^L	x_{i+1}^L	...	x_{N-1}^L	x_N^L

Fig.3 The left channel of audio M

The left channel is just predicted to select the smooth regions in the right channel. There is no message will be embedded in the left channel.

Usually, x_i^L's has strong local correlation with the context. So the local adjacent points $\{x_{i-k}^L \dots x_{i-3}^L, x_{i-2}^L, x_{i-1}^L, x_{i+1}^L, x_{i+2}^L, x_{i+3}^L \dots x_{i+k}^L\}$ can be used to predict x_i^L. The prediction value $\tilde{x}_i^L$ is given by (23):

$$\tilde{x}_i^L = \sum_{p=-k}^{-1} v_p x_{i-p}^L + \sum_{p=1}^{k} v_p x_{i-p}^L \quad (23)$$

Table.1 Correlation coefficient of ten audio clips

Index	Relevance
7	0. 9983
10	0. 7463
17	0. 8554
24	0. 8345
29	0. 9603
39	0. 5234
45	0. 6091
58	0. 7974
62	0. 7089
70	0. 9853

where v_p's are the prediction coefficients. In the prediction model, let $k = 3$, that means a sample point is estimated by the past three samples and the future three samples. The difference between predicted value $\tilde{x}_i^L$ and actual value x_i^L is calculated as:

$$e_i^L = \tilde{x}_i^L - x_i^L \quad (24)$$

We use the least squares regression method to get the best prediction coefficient $\mathbf{v_p}(p = -3, -2, -1, 1, 2, 3)$ such that

$$\mathbf{X_p} * \mathbf{v_p} = \mathbf{y_p} \quad (25)$$

where $\mathbf{X_p}$ is a 3×6 matrix

$$\mathbf{X_p} = \begin{bmatrix} \tilde{x}_{i-4}^L & \tilde{x}_{i-3}^L & \tilde{x}_{i-2}^L & \bar{x}_i^L & x_{i+1}^L & x_{i+2}^L \\ \tilde{x}_{i-3}^L & \tilde{x}_{i-2}^L & \tilde{x}_{i-1}^L & x_{i+1}^L & x_{i+2}^L & x_{i+3}^L \\ \tilde{x}_{i-2}^L & \tilde{x}_{i-1}^L & \bar{x}_i^L & x_{i+2}^L & x_{i+3}^L & x_{i+4}^L \end{bmatrix} \quad (26)$$

and $\mathbf{v_p} = [v_{-1} \quad v_{-2} \quad v_{-3} \quad v_1 \quad v_2 \quad v_3]'$, $\mathbf{y_p} = [\tilde{x}_{i-1}^L \quad \bar{x}_i^L \quad x_{i+1}^L]'$.

We use the approximate value $\bar{x}_i^L$ to replace x_i^L as shown in (27) so that the data can be lossless recovered at the receiver side

$$\bar{x}_i^L = (\tilde{x}_{i-1}^L + x_{i+1}^L)/2 \quad (27)$$

According to the least-square method, the best prediction coefficients are given by

$$\mathbf{v_p} = (\mathbf{X}'_\mathbf{p}\mathbf{X_p} + \mathbf{w})^{-1}\mathbf{X}'_\mathbf{p}\mathbf{y_p} \quad (28)$$

3.3 Payload assignment

We adaptively assign the payloads according to the degree of smoothness of the right channel, which is estimated by the information from the left channel. With the method described in Subsection prediction model in left channel , we can get the PEs of the left channel such that

$$E = (e_1^L, e_2^L, e_3^L, e_4^L, \dots, e_N^L) \quad (29)$$

We define the set of the smooth samples in the right channel as

$$S^R = \{x_i^R \ \mid \ |e_i^L| < tr\} \quad (30)$$

where *tr* is a threshold and we select *tr* as an integer. The set of S^R is then divided into a series of subsets according to the degrees of smoothness such that

$$l_j^R = \{x_i^R \ \mid j - 1 \le \ |e_i^L| < j\} \quad 1 \le j \le tr \qquad (31)$$

the set of S^R can be represented equally as

$$S^R = (l_1^R, l_2^R, l_3^R, \dots, l_j^R, \dots, l_{tr}^R) \quad (32)$$

Then, we use a randomly generated message to do a tentative embedding in each subset l_j^R using the basic method in single channel, by which we can estimate the capacity c_j and corresponding distortion d_j of the subset l_j^R. The c_j is the number of embeddable watermark bits, and d_j is calculated as (33).

$$d_j = \frac{\sum_{i=1}^{N}(\hat{x}_i^R - x_i^R)^2}{\sum_{i=1}^{N}(x_i^R)^2} \ , (33)$$

where $x_i^R \in l_j^R$, and $\hat{x}_i^R$ is the sample after tentative embedding. With such method, we get the estimated capacity in each subset denoted as

$$C^R = \ (c_1, c_2, c_3, \dots c_j, \dots, c_{tr}) \ , (34)$$

and the corresponding distortion set denoted as

$$D^R = \ (d_1, d_2, d_3, \dots d_j, \dots, d_{tr}) \ . (35)$$

For a given message length C, partial sample subsets are enough for accommodate the message. To get the optimal combination of subsets for C, we solve the following 0-1 programming problem.

$$\begin{cases} minimize \quad \sum_{j=1}^{tr} h_j * d_j \\ subject\ to\ \sum_{j=1}^{tr} h_j * c_j \ge C \end{cases} \quad h_j \in \{0,1\} \quad (36)$$

The optimal solution is denoted as

$$S_o^R = (l_{o1}^R, l_{o2}^R, \dots, l_{oh}^R) \quad (0 < o1 < o2 < \cdots < oh \le tr) \qquad (37)$$

Finally, with the basic method in single channel, we embed message into S_o^R in turn according to the location in the right of audio M.

3.4 Embedding procedure

There are some auxiliary information should be embedded into the cover audio (the right channel of audio M) for extraction and restoration. The parameters t, S_o^R, u_q^e and u_q^o are necessary to be embedded. The required space is calculated as follows.

1) Usually, $0 < t <15$ is enough for the method, and it needs 4 bits.
2) $S_o^R = (l_{o1}^R, l_{o2}^R, \dots, l_{oh}^R) \quad (0 < o1 < o2 < \cdots < oh \le tr)$. Usually $tr = 40$ is enough, and thus we use 40 bits to label which subset is chosed for embedding with the bit "1" represents selected subset.
3) u_q^e and u_q^o are decimals range from -1 to 1, and they need 30 bits.

The total size of the auxiliary information is 74 bits which occupies only a small amount of samples in audio M. To ensure adequate space, we use the last 80 samples in the right channel of audio M to embed the auxiliary information.

The details of embedding procedure are stated as follows.

Step1: Replace the LSBs of the last 80 samples in the right channel of audio M with parameters t, S_o^R, u_q^e and u_q^o, and then the LSBs will be reversibly embedded as part of the watermark message.

Step2: Select the optimal embedding samples S_o^R in the right channel of audio M according to the prediction error in the left channel of audio M.

Step3: For the selected samples S_o^R, we call the algorithm in Section 2 for embedding.

3.5 Extraction and restoration procedure

In the extracting procedure, our purpose is to get watermark message from the marked audio and restore the original audio. The details are stated as follows:

Step1: Preprocess the marked audio. Read the last 80 samples in right channel to get threshold t in (17), chosen region S_o^R in (37), coefficients u_q^e and u_q^o.

Step2: Calculate the prediction value of left channel and the prediction error $E = (e_1^L, e_2^L, e_3^L, e_4^L, \dots, e_N^L)$. Determine the embedding region according to S_o^R.

Step3: Recover $\hat{e}_i^R$ of right channel according to u_q^e and u_q^o, first odd set, then even set . Extract $watermark$ and the LSB using (19) and restore the audio M' using (21).

Step4: Use the LSB to replace the last 80 samples in right channel to get original audio M.

4. EXPERIMENTAL RESULTS

In the testing, we get 4 audio files for example in audio database [37] to evaluate the performance of the proposed algorithm. Uniformly, all of the clips are standard stereo audio and the sampling frequency is 44.1k. We intercept 200000 audio points for the audio files in order to perform intuitively. Embedding capacity and audio distortion are two important criterions to be calculated. The embedding capacity is represented by the amount of data embedded in audio. And the signal to noise ratio (SNR) is used to measure the watermark distortion.

$$\mathrm{SNR} = 10\lg\left(\frac{\sum_{i=1}^{N}\left(\hat{x}_i^R - x_i^R\right)^2}{\sum_{i=1}^{N}(x_i^R)^2}\right) \quad (38)$$

We compare the proposed algorithm with three existing work in reversible audio data hiding algorithm [29], [35], [36] for clip 08, 16, 53, 70. *Fig.4, 5, 6, 7* show the experimental performance of four algorithm on the example clips.

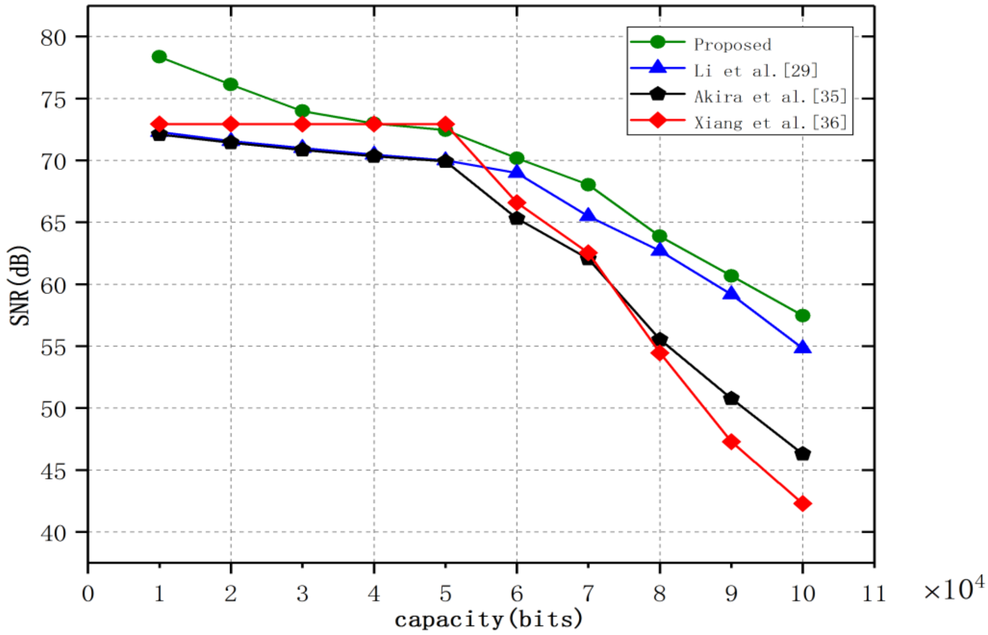

Fig4. Distortion comparison for clip 08

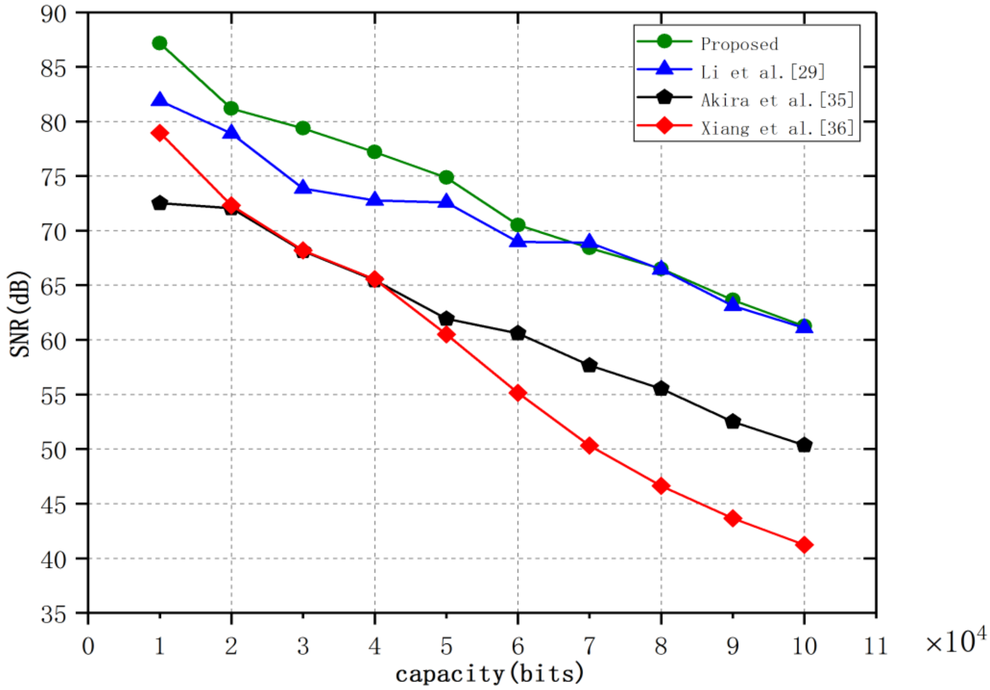

Fig5. Distortion comparison for clip 16

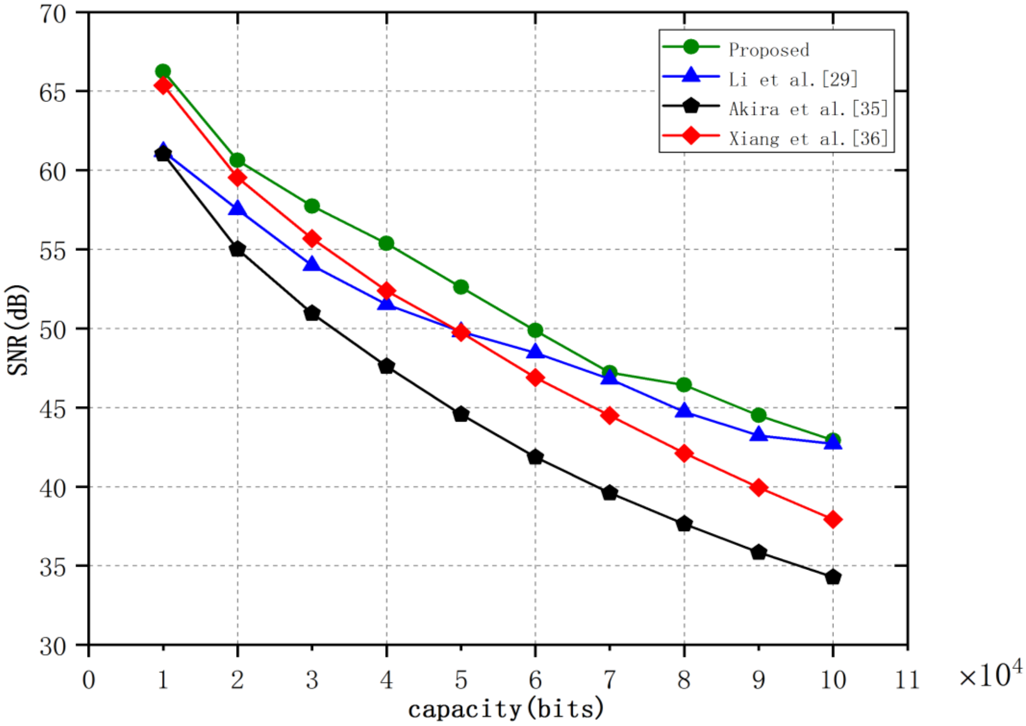

Fig6. Distortion comparison for clip 53

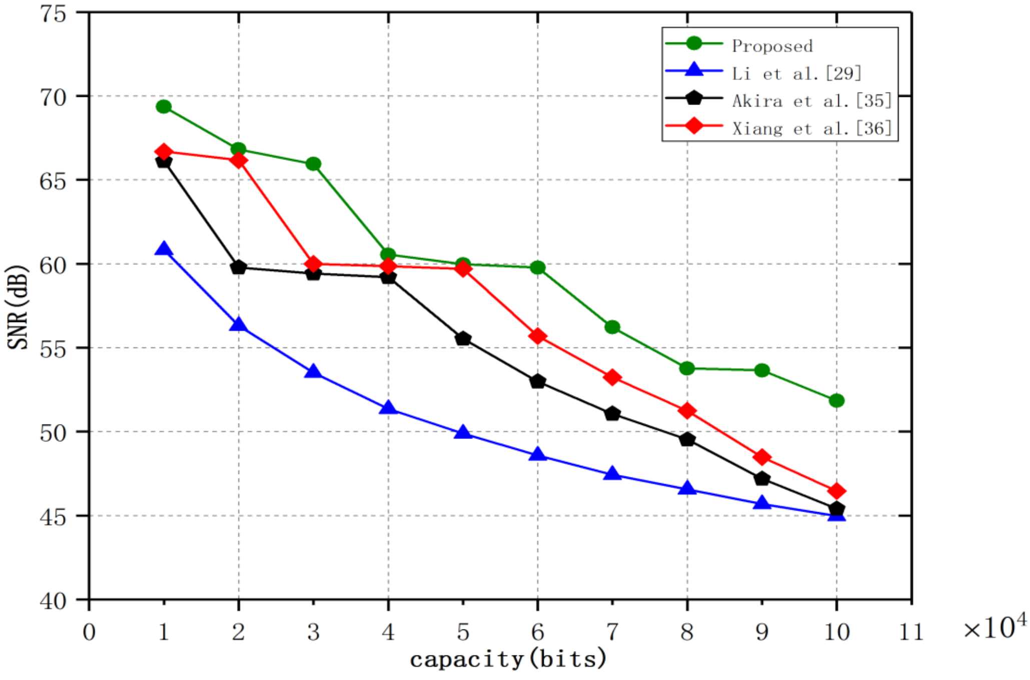

Fig7. Distortion comparison for clip 70

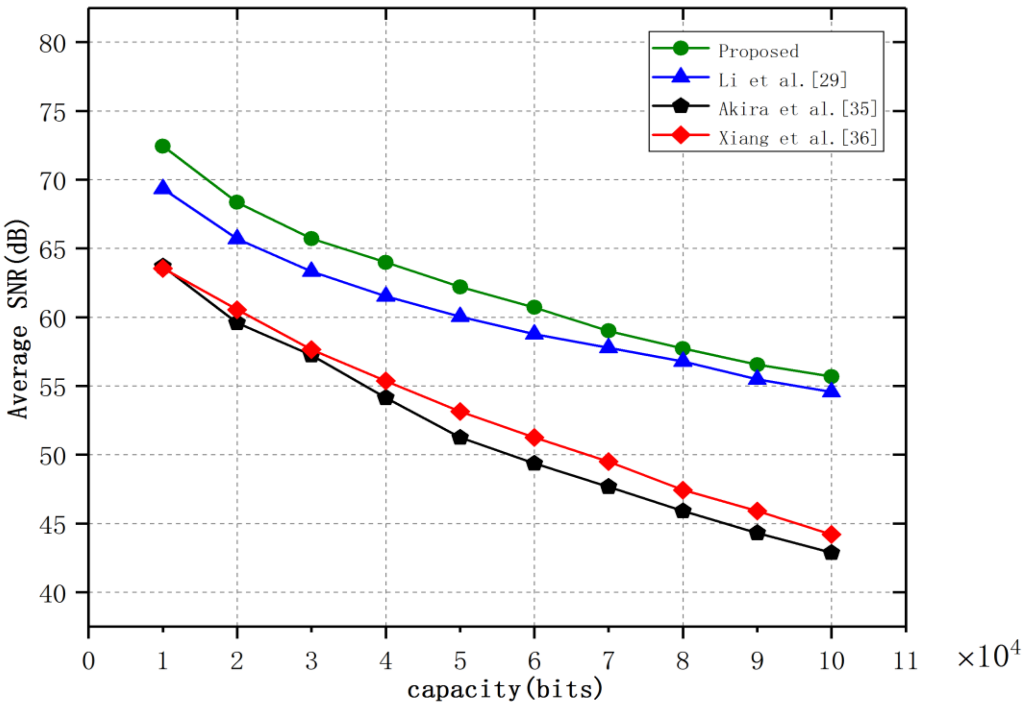

Fig8. Average SNR for stereo audio database for different capacity

Fig.8 shows the average SNR of 70 audio clips in database [37] for different capacity. The result of these methods are listed in *Table 2*. We can observe that the proposed method outperforms the existing work in reversible audio data hiding [29], [35], [36].

Table 2. Average SNR for stereo audio database [37] for different capacity

Capacity×10^4	1	2	3	4	5	6	7	8	9	10
Proposed	72.45	68.35	65.72	63.98	62.20	60.71	59.01	57.73	56.55	55.68
Li et al.[29]	69.35	65.69	63.34	61.52	60.04	58.77	57.79	56.78	55.50	54.58
Akira et al.[35]	63.69	59.60	57.24	54.14	51.25	49.37	47.67	45.90	44.31	42.88
Xiang et al.[36]	63.55	60.55	57.65	55.36	53.14	51.25	49.50	47.44	45.90	44.21

5. CONCLUSION

In this paper, we proposed a reversible watermarking algorithm for stereo audio based on the inter-channel correlation. The message is only embedded into the right channel while the embedding regions are determined with the help of the smoothness degrees in the left channel. By exploring such inter-channel correlation, we can effectively avoid region that may introduce large modification costs. Experimental results illustrate that the proposed method achieves lower distortion than several traditional methods that use single audio channel.

6. ACKNOWLEDGEMENTS

This work was supported in part by the Natural Science Foundation of China under Grant U1636201, 61572452.

REFERENCES

[1] H.S. Malvar, & D.A.F.Florencio. (2003). Improved spread spectrum: a new modulation technique for robust watermarking. IEEE Transactions on Signal Processing, 51(4), 898-905.

[2] Yashar Naderahmadian, & Saied Hosseini-Khayat. (2014). Fast and robust watermarking in still images based on QR decomposition. Multimed Tools Applications, 72(3), 2597–2618.

[3] Huawei Tian, Yanhui Xiao, & Gang Cao. (2016). Robust watermarking of mobile video resistant against barrel distortion. China Communications, 13(9), 131-138.

[4] Asha Rani, Balasubramanian Raman, & Sanjeev Kumar. (2014). A robust watermarking scheme exploiting balanced neural tree for rightful ownership protection. Multimed Tools Applications, 72(3), 2225–2248.

[5] Xinpeng Zhang, & Shuozhong Wang. (2009). Fragile watermarking scheme using a hierarchical mechanism. Signal processing, 89(4), 675-679.

[6] Xinpeng Zhang, & Shuozhong Wang. (2008). Fragile Watermarking With Error-Free Restoration Capability. IEEE Transactions on Multimedia, 10(8), 1490-1499.

[7] Chuan Qin, Chin-Chen Chang, & Pei-Yu Chen. (2012). Self-embedding fragile watermarking with restoration capability based on adaptive bit allocation mechanism. Signal processing, 92(4), 1137-1150.

[8] J. M. Barton. (1997). Method and apparatus for embedding authentication information within digital data, US Patent 5.

[9] M. U. Celik, G. Sharma, A. M. Tekalp, & E. Saber. (2005). Lossless generalized-LSB data embedding. IEEE Trans. Image Process, 14(2), 253–266.

[10] S. J. Lin, & W. H. Chung. (2011). The scalar scheme for reversible information-embedding in gray-scale signals: capacity evaluation and code constructions. IEEE Trans. Inf. Forens. Security, 7(4), 1155–1167.

[11] W. Zhang, B. Chen, & N. Yu. (2012). Improving various reversible data hiding schemes via optimal codes for binary covers. IEEE Trans. Image Process, 21(6), 2991–3003.

[12] J. Tian. (2003). Reversible data embedding using a difference expansion. IEEE Trans. Circuits Syst. Video Technol, 13(8), 890–896.

[13] Z. Ni, Y. Q. Shi, N. Ansari, & W. Su. (2006). Reversible data hiding. IEEE Trans. Circuits Syst. Video Technol, 16(3), 354–362.

[14] D. M. Thodi, & J. J. Rodriguez. (2007). Expansion embedding techniques for reversible watermarking. IEEE Trans. Image Process, 16(3), 721–730.

[15] W. Hong, T. S. Chen, & C. W. Shiu. (2009). Reversible data hiding for high quality images using modification of prediction errors. J. Syst. Softw, 82(11), 1833–1842.

[16] J. Wang, & J. Ni. (2013). A GA optimization approach to HS based multiple reversible data hiding. In Proc. IEEE WIFS, 203–208.

[17] B. Ou, X. Li, Y. Zhao, R. Ni, & Y. Q. Shi. (2013). Pairwise prediction-error expansion for efficient reversible data hiding. IEEE Trans. Image Process, 22(12), 5010–5021.

[18] X. Li, B. Yang, & T. Zeng. (2011). Efficient reversible watermarking based on adaptive prediction-error expansion and pixel selection. IEEE Trans. Image Process, 20(12), 3524–3533.
[19] L. Kamstra, & H. J. A. M. Heijmans. (2005). Reversible data embedding into images using wavelet techniques and sorting. IEEE Trans. Image Process, 14(12), 2082–2090.
[20] X. Li, J. Li, B. Li, & B. Yang. (2013). High-fidelity reversible data hiding scheme based on pixel-value-ordering and prediction-error expansion. Signal Process, 93(1), 198–205.
[21] A. M. Alattar. (2004). Reversible watermark using the difference expansion of a generalized integer transform. IEEE Trans. Image Process 13(8), 1147–1156.
[22] X. Chen, X. Li, B. Yang, & Y. Tang. (2010). Reversible image watermarking based on a generalized integer transform. In Proc. IEEE ICASSP, 2382–2385.
[23] F. Peng, X. Li, & B. Yang. (2012). Adaptive reversible data hiding scheme based on integer transform. Signal Process, 92(1), 54–62.
[24] VDV. Michiel, V.L. Arno, & B. Fons. (2003). Reversible Audio Watermarking. Audio Engineering Society, 5818.
[25] Yan D, & Wang R. (2008). Reversible Data Hiding for Audio Based on Prediction Error Expansion. Intelligent Information Hiding and Multimedia Signal Processing, 249–252.
[26] J.J.Garcia-Hernandez. (2012). Exploring reversible digital watermarking in audio signals using additive interpolation-error expansion. Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP), 2012 Eighth International Conference on, 40.
[27] Bradley B, & Alattar A. (2015). High-capacity invertible data-hiding algorithm for digital audio. SPIE, 789.
[28] Fei Wang, Zhaoxin Xie, & Zuo Chen. (2014). High Capacity Reversible Watermarking for Audio by Histogram Shifting and Predicted Error Expansion. The Scientific Word Journal.
[29] Shijun Xiang, & Zihao Li. (2017). Reversible audio data hiding algorithm using noncausal prediction of alterable orders. EURASIP Journal on Audio, Speech, and Music Processing, 4.
[30] Mingyu Li, Yuhua Jiao, & Xiamu Niu. (2008). Reversible Watermarking for Compressed Speech. Intelligent Systems Design and Applications, 2008. ISDA '08. Eighth International Conference on, 197–201.
[31] Chen OTC, & Liu CH. (2007). Content-Dependent Watermarking Scheme in Compressed Speech With Identifying Manner and Location of Attacks. IEEE Transactions on Audio, Speech, and Language Processing, 1605–1616.
[32] Xuping Huang, Nobutaka Ono, Isao Echizen, & Akira Nishimura. (2013). Reversible Audio Information Hiding Based on Integer DCT Coefficients with Adaptive Hiding Locations. IWDW, 376-389.
[33] Yan Yang, Rong Huang, & Mintao Xu. (2009). A Novel Audio Watermarking Algorithm for Copyright Protection Based on DCT Domain. Electronic Commerce and Security, 2009. ISECS '09. Second International Symposium on, 184–188.

[34] Masashi Unoki, & Ryota Miyauchi. (2011). Reversible Watermarking for Digital Audio Based on Cochlear Delay Characteristics. Intelligent Information Hiding and Multimedia Signal Processing.
[35] Akira Nishimura. (2011). Reversible audio data hiding using linear prediction and error expansion. Proc. of IIHMSP2011, 318–321.
[36] Shijun Xiang. (2012). Non-integer expansion embedding for prediction-based reversible watermarking. Proc. 14th Int. Conf, 224–239.
[37] EBU Committee: sound quality assessment material recordings for subjective tests [online]. Available: https://tech.ebu.ch/publications/sqamcd.