IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received July 18, 2020, accepted August 6, 2020, date of publication August 11, 2020, date of current version September 17, 2020.

*Digital Object Identifier 10.1109/ACCESS.2020.3015851*

# A High-Capacity Reversible Data Hiding Scheme Using Dual-Channel Audio

**HENG YU, RANGDING WANG, (Member, IEEE), LI DONG, (Member, IEEE), DIQUN YAN, (Member, IEEE), YONGKANG GONG, AND YUZHEN LIN**
Faculty of Electrical Engineering and Computer Science, Ningbo University, Ningbo 315211, China

Corresponding author: Rangding Wang (wangrangding@nbu.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant U1736215, Grant 61672302, and Grant 61901237; in part by the Zhejiang Natural Science Foundation under Grant LY20F020010 and Grant LY17F020010; and in part by K. C. Wong Magna Fund in Ningbo University.

**ABSTRACT** In recent years, the reversible data hiding (RDH) based on dual stego cover is developing rapidly because of its high capacity and low distortion. For image case, however, two consecutive images of the same image will draw the attention of adversaries during transmission. In this article, we propose a high-capacity RDH scheme using dual-channel audio, by exploiting the natural dual-channel property of the stereo audio. Specifically, we first convert secret message into novenary digits, which could increase the embedding capacity. Then, the magic matrix is used to embed the secret digit into a single-channel audio to generate two single-channel stego-audio. Finally, the two single-channel stego-audio is combined as a convention dual-channel audio. Extensive experiments have demonstrated that our proposed method could significantly boost the stego quality (the SNR is improved by 16% on average), when comparing with the state-of-the-art methods.

**INDEX TERMS** Reversible data hiding, magic matrix, audio, dual channel.

## I. INTRODUCTION

This Data hiding is a technique of embedding secret messages in digital media, which exploits the redundancy of human perception to achieve covert communication. There are several common schemes of traditional steganography: LSB [1], F5 [2], STC [3]. However, these schemes cause permanent distortion of the original cover which is not allowed in some critical scenarios [4]. To protect these signals, reversible data hiding (RDH) [5] is developed. In addition to covert communication, RDH can make the cover erasable so that the storage space can be used repeatedly. Furthermore, there are some other applications [6]: reversible adversarial example, reversible visual transformation and reversible image processing.

The existing RDH schemes can be roughly divided into the following four categories: difference expansion (DE), histogram shifting (HS), pixel value ordering (PVO) and dual-image.

Difference expansion (DE)-based: difference expansion (DE) [7], [8] was first proposed by Tian *et al.*. The latter prediction error expansion (PEE) [9], [10] are based on DE. The core of these methods is to first generate a small prediction-error through the correlation between adjacent pixels, then convert this prediction-error into binary format, and finally embed the secret message bit in the least significant bit of the binary prediction-error in an expansion way. The strength of the prediction mechanism directly affects the performance of such algorithms. Therefore, the incurred distortion of these methods is usually large.

Histogram shifting (HS)-based: histogram shifting (HS) [11], difference histogram shifting (DHS) [12] and prediction error histogram shifting (PEHS) [13]. This type of scheme first generates a histogram based on pixels or pixel differences or prediction errors, and then finds the peaks and zeros in the histogram, the values of both sides of the peak point and zero to move one bit to make room. Finally, the embedding is performed by moving the value between the peak point and the zero point. Because the embedding is by moving the value between peak point and zero point, the embedding capacity is often limited.

Pixel value ordering (PVO)-based: Li *et al.* [14] presented a blocked-based RDH method by combining PEE and pixel value ordering (PVO). PVO is kept unaltered by modifying

The associate editor coordinating the review of this manuscript and approving it for publication was Jingchang Huang.

the maximum in a block increase at most one while the minimum decrease at most one. The distortion thus can be lowered. But it only use the pixel where the prediction error is equal to one, so the embedding capacity is also lower. Hence, Peng *et al.* [15] proposed an improved PVO (IPVO) method. They use the pixel where the prediction error is equal to zero or one for embedding. It increase the embedding capacity to some extent and many subsequent research [16], [17] are based on them. Low embedded capacity is always an imperfection of these methods.

RDH based on dual images: owing to the above traditional methods are faced with the problems of low embedded capacity or high distortion. Chang *et al.* [18] proposed a RDH method with dual images that implemented large capacity and guaranteed good image quality. They proposed a new information hiding scheme, in which two base-5 secret digits are embedded individually into two stego-images by using a specifically-designed magic matrix. Their scheme can achieve a considerable embedding rate (ER) of about 1.0 bit per pixel (bpp), and it provides good quality of stego-images that have an average PSNR value of 45 dB. Due to one quinary digit can represent secret message of two bits, while novenary can represent secret message of three bits. In 2013, The quinary representation of data bits was replaced in [19] with a novenary representation. This increased the embedding capacity to about 1.5bpp. Meanwhile, the quality of two stego-images remain around 39dB. In 2015, Qin *et al.* [20] proposed an RDH scheme based on exploiting modification direction (EMD) to generate two stego-images. The algorithm achieved an embedding rate that slightly above 1 bpp. However, the quality of the two stego-images was asymmetric. In addition, there are some works [21], [22] on further increase the embedding capacity.

However, most of the aforementioned methods are proposed for image case, or for the single-channel audio [10], [23]. Only a few works considered the audio for the dual-channel stego cover. We propose a novel RDH scheme for dual-channel audio in this article. First convert the binary secret message into the novenary digits, and then create a two-dimensional magic matrix made up of the numbers 0 to 8. We take a single-channel audio as the cover and use a sample pair to conceal a secret digit to generate a dual-channel stego-audio. Because a secret digit can represent the secret message of three bits and the secret message is all hidden in the left channel, the embedding rate can be up to 1.5 bit per sample (bps). Compared with existing schemes, our method achieve lower distortion.

The rest of the paper is organized as follows: Section II introduces the related work; Section III is our proposed method, including the embedding algorithm and extraction algorithm; Section IV is the experimental result and analysis; Section V concludes this work.

## II. RELATED WORKS

In this section, we introduce one typical for reversible data hiding based on dual images [19].

**TABLE 1. The pixel pair modification rules of Chang *et al.* [19].**

| $M, V$ | $x_1'$ | $x_2'$ |
|---|---|---|
| $M = V$ | $x_1$ | $x_2$ |
| $M = (V-2) \bmod 9$ | $x_1 - 1$ | $x_2 + 1$ |
| $M = (V+5) \bmod 9$ | $x_1 - 2$ | $x_2 + 2$ |
| $M = (V+3) \bmod 9$ | $x_1 - 3$ | $x_2 + 3$ |
| $M = (V+1) \bmod 9$ | $x_1 - 4$ | $x_2 + 4$ |
| $M = (V+2) \bmod 9$ | $x_1 + 1$ | $x_2 - 1$ |
| $M = (V+4) \bmod 9$ | $x_1 + 2$ | $x_2 - 2$ |
| $M = (V+6) \bmod 9$ | $x_1 + 3$ | $x_2 - 3$ |
| $M = (V-1) \bmod 9$ | $x_1 + 4$ | $x_2 - 4$ |

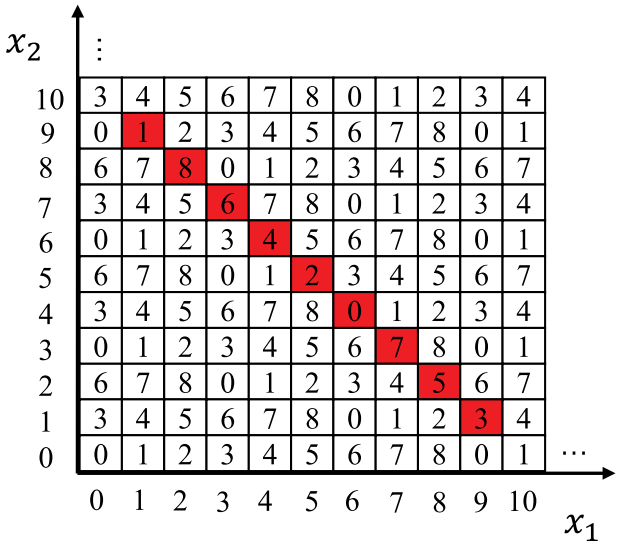


**FIGURE 1. A portion of the magic matrix, and the red area is the scope of the current pixel pair (5, 5) can modify.**

Chang *et al.* [19] developed the dual-image reversible data hiding method based on the magic matrix. The magic matrix is generated by $M(x_1, x_2) = (x_1 + 3 \times x_2) \bmod 9$, where the parameters, $x_1$ and $x_2$ denote two pixels ranged in [0, 255], and $M(x_1, x_2)$ represents the secret data that is shown in Fig. 1. The embedding algorithm is described below.

First, extract $n$ secret bits from the secret data $S$ and transform these extracted bits into a decimal value $d$, where the initial value of $n$ is set to be 4. If the decimal value $d$ is not equal to 8, then $n$–1. Then, extract one pixel $x_1$ from the cover image $I$ and copy it by $x_2 = x_1$. Finally, embed the novenary digit $d$ into the pixel pair $(x_1, x_2)$ by Table 1.
where $M = (x_1 + 3 \times x_2) \bmod 9$, and $x_1'$ represents the pixel of the first stego image and $x_2'$ denotes the pixel of the second stego-image.
The extraction and recovery algorithm is described below.

First, extract two pixels $x_1'$ and $x_2'$ from two stego-images. Next, obtain the novenary digit $d$ by $d = (x_1' + 3 \times x_2') \bmod 9$. Then, if the novenary digit $d$ is higher than 7, it is transformed into four secret bits. Otherwise, it is transformed into three secret bits. Finally, calculate the integer average value of two

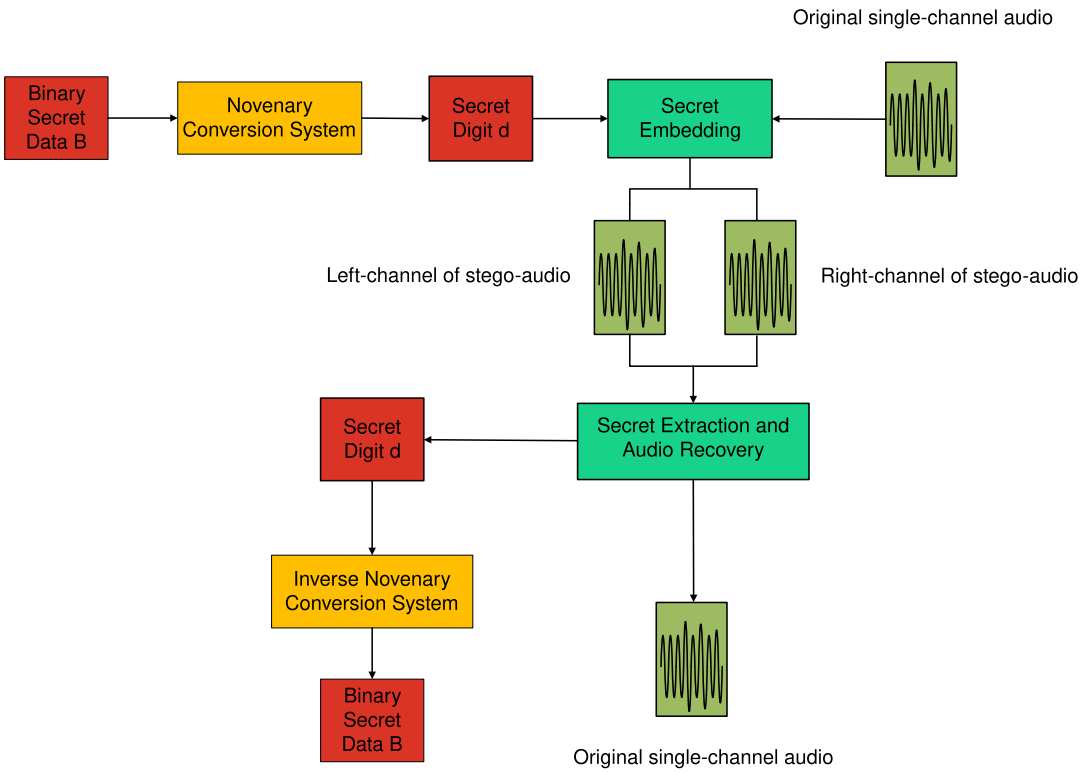


**FIGURE 2.** Flowchart of the proposed method.

stego-pixels $x_1'$ and $x_2'$ to derive the original pixel $x_1$, where $x_1 = \lfloor (x_1' + x_2') / 2 \rfloor$.

As one can see, when the pixel value is equal to $\{0, 1, 2, 3, 252, 253, 254, 255\}$, the pixel may become negative after the embedding operation. Therefore, the embedded capacity will be influenced. In addition, due to the current pixel pair is replaced by one of its four upperleft or bottom-right pixel pairs, in which its corresponding value is equivalent to secret data. In fact, the distortion caused by this method is relatively large. Therefore, there is still room for improvement.

## III. RDH OF DUAL-CHANNEL AUDIO

In this section, we propose a new reversible mechanism based on audio using magic matrix. The proposed algorithm takes single-channel audio as the cover and generate a dual-channel audio after embedding. The complete flow chart of the algorithm is shown in Fig. 2. Section III-A introduces the embedding algorithm. Section III-B introduces the extraction and recovery algorithm.

### A. EMBEDDING ALGORITHM

In order to guarantee the embedding capacity, we convert the binary secret data into novenary digits. This is because a novenary digit can represent at most $\log_2 9$ binary bits which can significantly increase the embedding capacity.

To ease of calculation, we adjust the range of the samples between [0, 65535] by adding 32768 to each sample. Before embedding, we create a magic matrix $M$ of size $65535 \times 65535$ by the following:

$$M(x_1, x_2) = (x_1 + 3 \times x_2) \; mod \; 9, \tag{1}$$

where $x_1, x_2 \in [0, 65535]$ are represent samples of audio, the magic matrix $M$ is shown in Fig. 3.

From Fig. 3, we can obviously see that any grid of size $3 \times 3$ include 0 to 8. We use a pair of sample to conceal a novenary secret digit. When the the corresponding value $M(x_1, x_2)$ of the current sample pair $(x_1, x_2)$ is not equal to the secret digit $d$, the current sample pair is replaced by one of its eight surrounding sample pairs $(t_1, t_2)$, in which its corresponding value is equivalent to the secret digit $d$. After the above operation, the left-channel audio $X_l'$ is obtained by following:

$$X_l' = (t_1, t_2), \tag{2}$$

When the corresponding value $M(x_1, x_2)$ of the current sample pair $(x_1, x_2)$ is equal to the secret digit $d$, the left-channel audio is obtained by following:

$$X_l' = (x_1, x_2). \tag{3}$$

To ensure invertibility, we calculate $c$ by

$$c = ((M(x_1, x_2) + d) \, mod \, 9), \tag{4}$$

Clearly, $c$ is between 0 and 8. We find a sample pair $(h_1, h_2)$ around the sample pair $(x_1, x_2)$ which the corresponding value

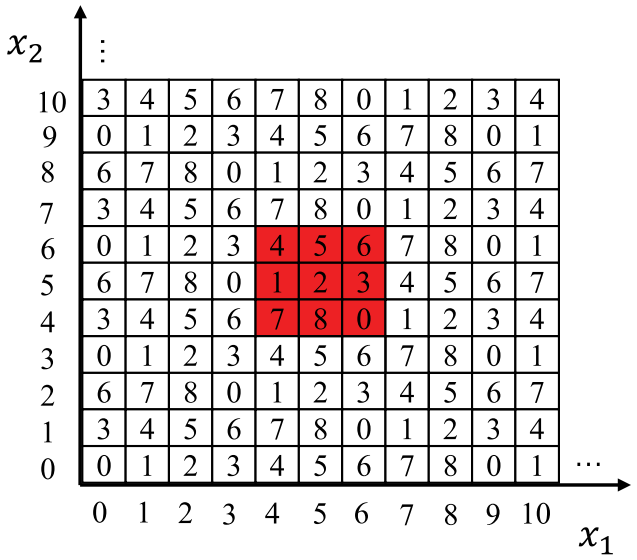


**FIGURE 3.** A portion of the magic matrix, and the red area is the scope of the current sample pair (5, 5) can modify.

$M(h_1, h_2)$ is equal to $c$. Then, the right-channel audio is obtained by following:

$$X'_r = (h_1, h_2). \tag{5}$$

### B. EXTRACTION AND RECOVERY ALGORITHM

After receiving the dual-channel stego-audio, we use the left channel to extract the secret digits. First, extract a sample pair $(x_1^l, x_2^l)$ from the left-channel audio $X'_l$. Then, the corresponding value $M(x_1^l, x_2^l)$ is the secret digit $d$.

Upon restoring the original audio, we extract a sample pair $(x_1^r, x_2^r)$ from the right-channel audio $X'_r$ and obtain $c = M(x_1^r, x_2^r)$. If $c$ is more than equal to $d$, we find the original sample pair $(x_1, x_2)$ around $(x_1^l, x_2^l)$ where $M(x_1, x_2) = c - d$. Otherwise the original sample pair is $(x_1, x_2)$ around $(x_1^l, x_2^l)$ where $M(x_1, x_2) = c + 9 - d$.

In our method, the modification of one sample of the audio is at most two. For example, suppose the binary secret data is $(101101)_2$ that the corresponding novenary secret digits is $(50)_9$. The digit $d = 5$ is embedded first, suppose the original audio $X$ is (5, 5). Find $M(5, 6)$ is equal to 5 around the corresponding value of the sample pair $(x_1, x_2)$, then the left-channel stego-audio $X'_l$ is (5, 6). Calculate the $c = (M(5, 5) + 5) \, mod \, 9 = 7$, because $M(4, 4) = 7$, so the right-channel stego-audio $X'_r$ is (4, 4). When extracting the secret data, extract the sample pair (5, 6) from left-channel audio $X'_l$. The secret digit is $d = M(5, 6) = 5$. Then extract the sample pair (4, 4) from right-channel audio $X'_r$, $c = M(4, 4) = 7$, because $c > d$ and $M(5, 5) = 7 - 5 = 2$, original audio $X$ is (5, 5). The second digit $d = 0$ is the same operation.

## IV. EXPERIMENTAL RESULTS

In this section, the proposed method is compared with Chang *et al.* [19] and Xiang *et al.* [10].

**Algorithm 1** Embedding Algorithm

Variable $N$ denotes the length of audio $X$, $B$ denotes the binary secret data, $D$ denotes the novenary secret digits, $D_i$ denotes the $i$th secret digit.
**Input:** Single-channel audio $X$ and binary secret data $B$
**Output:** Left-channel stego-audio $X'_l$, Right-channel stego-audio $X'_r$
Convert binary secret data $B$ into the novenary digits $D$
**for** $j = 1 : 2 : N - 1, i = 1 : floor(N/2)$ **do**
  $a = X(j) + 32768$
  $b = X(j + 1) + 32768$
  **if** $(a == 0) \,||\, (a == 65535) \,||\, (b == 0) \,||\, (b == 65535)$ **then**
    $X'_l(j) = X'_r(j) = a - 32768$
    $X'_l(j + 1) = X'_r(j + 1) = b - 32768$
  **end if**
  **for** $k = 1 : 9$ **do**
    Find a sample pair $(a_k, b_k)$ around $(a, b)$
    **if** $M(a_k, b_k) == D_i$ **then**
      $X'_l(j) = a_k - 32768$
      $X'_l(j + 1) = b_k - 32768$
    **end if**
  **end for**
  $e = (M(a, b) + D_i) \bmod 9$
  **for** $k = 1 : 9$ **do**
    Find a sample pair $(a_k, b_k)$ around $(a, b)$
    **if** $e == M(a_k, b_k)$ **then**
      $X'_r(j) = a_k - 32768$
      $X'_r(j + 1) = b_k - 32768$
    **end if**
  **end for**
**end for**
**return** $X'_l, X'_r$

By comparing Fig. 4 and Fig. 5, it's not difficult to find that the audio waveform of proposed method is similar to original audio. It's because the proposed method modifies a single sample value by one at most. Fig. 6 illustrates this result. Similarly, Fig. 7 shows the difference between the original audio and steganographic audio of Chang's method. It can be found that the maximum modification amplitude of Chang's method to a single sample is reached to four. In contrast to the above, Xiang is a method of prediction error expansion. Due to the size of the prediction error is not

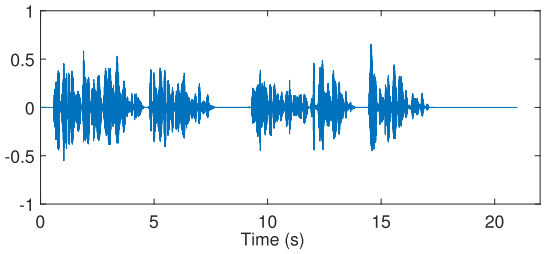


**FIGURE 4.** The original audio waveform.

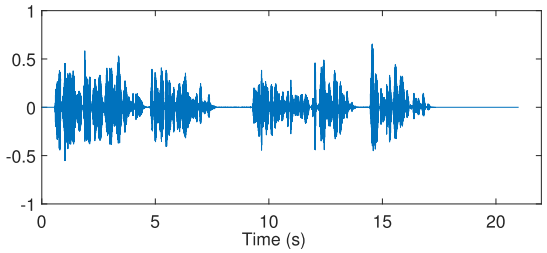


**FIGURE 5.** The audio waveform of proposed method.

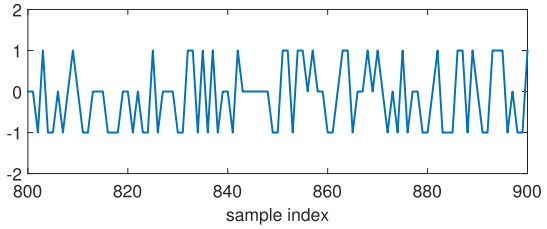


**FIGURE 6.** A portion of the difference between the original audio and the audio of proposed method.

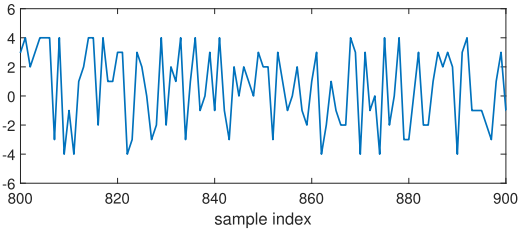


**FIGURE 7.** A portion of the difference between the original audio and the audio of Chang's method.

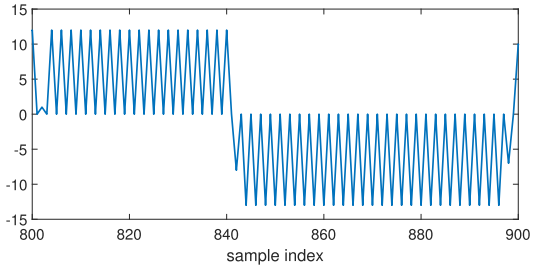


**FIGURE 8.** A portion of the difference between the original audio and the audio of Xiang's method.

controllable, the distortion will be severe after the expansion operation. It can be seen from Fig. 8 that the distortion caused by Xiang's method is obviously larger. In fact, there is no theoretical upper limit to this distortion.

To better describe the algorithm distortion, we use mathematical expectation(**ME**) to quantify it. As shown in Table 2, in the case of embedding one secret digit, the probability of Chang's method modifying one sample by 0, 1, 2, 3, 4 are $\frac{1}{9}$, $\frac{2}{9}$, $\frac{2}{9}$, $\frac{2}{9}$, $\frac{2}{9}$, respectively. Obviously the mathematical expectation is equal to $\frac{20}{9}$. However, for our method, the maximum modification of one sample is one, so the modifier of two to four's probability is zero. When the amount of modification

**TABLE 2.** The probabilities and mathematical expectations(ME) for different modifications.

| Method | 0 | 1 | 2 | 3 | 4 | ME |
|---|---|---|---|---|---|---|
| Chang [19] | $\frac{1}{9}$ | $\frac{2}{9}$ | $\frac{2}{9}$ | $\frac{2}{9}$ | $\frac{2}{9}$ | $\frac{20}{9}$ |
| Proposed | $\frac{1}{9}$ | $\frac{8}{9}$ | 0 | 0 | 0 | $\frac{8}{9}$ |

is 0, 1, the corresponding probabilities are $\frac{1}{9}$, $\frac{8}{9}$, respectively. Obviously the mathematical expectation is equal to $\frac{8}{9}$. Clearly $\frac{8}{9}$ is less than $\frac{20}{9}$, which means that our method has a lower distortion than Chang's method.

The performance of the three approaches is evaluated on 70 standard audio files with a sampling rate of 44.1 kHz. The test set is available at [24] and comprises: alignment signals (clips 1 and 2), artificial signals (3 to 7), single instruments (8 to 43), vocal (44 to 48), speech (49 to 54), solo instruments (55 to 60), vocal with orchestra (61 to 64), orchestra (65 to 68) and pop music (69 and 70). Set these audio to single-channel during the experiment.

**Algorithm 2** Extraction and Recovery Algorithm

Variable $N$ denotes the length of audio $X$, $D_i$ denotes the $i$th secret digit.
**Input:** Dual-channel audio $X'$
**Output:** Single-channel auido $X$, binary secret data $B$
**for** $j = 1 : 2 : N-1, i = 1 : floor(N/2)$ **do**
  $x_1^l = X_l'(j) + 32768$
  $x_2^l = X_l'(j+1) + 32768$
  $D_i = M(a_l, b_l)$
  $x_1^r = X_r'(j) + 32768$
  $x_2^r = X_r'(j+1) + 32768$
  $c = M(x_1^r, x_2^r)$
  **if** $c \geq d$ **then**
    **for** $k = 1 : 9$ **do**
      Find a sample pair $(a_k, b_k)$ around $(x_1^l, x_2^l)$
      **if** $c - d == M(a_k, b_k)$ **then**
        $X(j) = a_k - 32768$
        $X(j+1) = b_k - 32768$
      **end if**
    **end for**
  **else if** $c < d$ **then**
    **for** $k = 1 : 9$ **do**
      Find a sample pair $(a_k, b_k)$ around $(x_1^l, x_2^l)$
      **if** $c + 9 - D_i == M(a_k, b_k)$ **then**
        $X(j) = a_k - 32768$
        $X(j+1) = b_k - 32768$
      **end if**
    **end for**
  **end if**
**end for**
Convert the novenary digits $D$ into the binary secret data $B$
**return** $X, B$

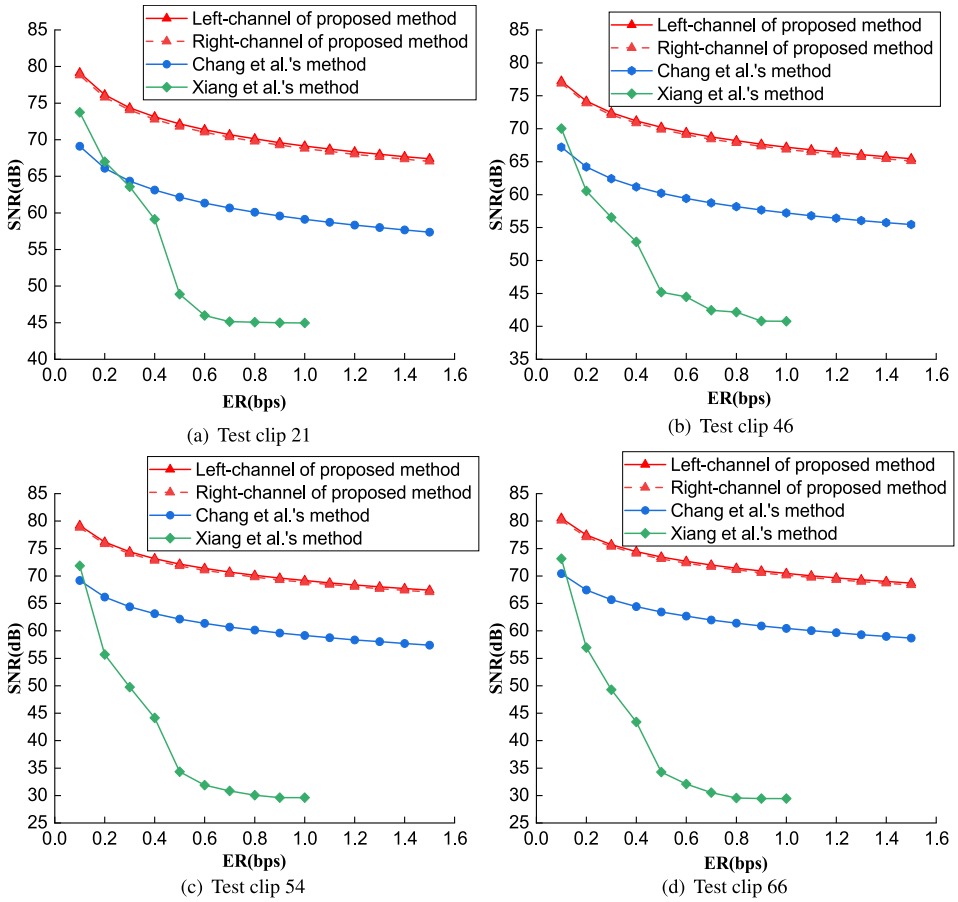


**FIGURE 9.** SNR value for proposed method, Xiang [10] and Chang [19] on single instruments (clip 21), vocal (clip 46), speech (clip 54) and orchestra (clip 66). The left channel and right channel audio generated by our method is very similar (the two red lines are almost overlapped).

**TABLE 3.** SNR results from proposed method, Xiang [10] and Chang [19] of nine kinds of audio (ER = 1bps).

| Audio | Xiang [10] | Chang [19] | Proposed | |
|---|---|---|---|---|
| Alignment signals | 36.46 | 56.34 | 65.36 | 65.03 |
| Artificial signals | 46.37 | 70.07 | 81.28 | 81 |
| Single instruments | 40.53 | 62.05 | 71.98 | 71.65 |
| Vocal | 38.74 | 60.57 | 70.29 | 69.96 |
| Speech | 42.65 | 61.71 | 71.58 | 71.25 |
| Solo instruments | 45.28 | 60.31 | 69.96 | 69.63 |
| Vocal with orchestra | 40.68 | 63.02 | 73.1 | 72.77 |
| Orchestra | 34.87 | 58.48 | 67.84 | 67.51 |
| Pop music | 24.45 | 59.26 | 68.74 | 68.41 |
| **Average** | 38.89 | 61.31 | 71.13 | 70.8 |

**TABLE 4.** ODG results from proposed method, Xiang [10] and Chang [19] of nine kinds of audio (ER = 1bps).

| Audio | Xiang [10] | Chang [19] | Proposed | |
|---|---|---|---|---|
| Alignment signals | −0.476 | −0.144 | −0.144 | −0.156 |
| Artificial signals | −0.856 | −2.02 | −0.359 | −0.403 |
| Single instruments | −0.374 | −0.744 | −0.067 | −0.079 |
| Vocal | −0.976 | −0.368 | −0.006 | −0.007 |
| Speech | −1.574 | −0.2 | −0.001 | −0.001 |
| Solo instruments | −1.2 | −0.201 | −0.041 | −0.047 |
| Vocal with orchestra | −1.208 | −0.032 | −0.019 | −0.018 |
| Orchestra | −0.805 | −0.638 | −0.022 | −0.027 |
| Pop music | −3.913 | −0.104 | −0.006 | −0.007 |
| **Average** | −1.265 | −0.495 | −0.074 | −0.083 |

We use two common metrics, signal-to-noise-ratio (**SNR**) and objective difference grade (**ODG**) to evaluate the distortion of stego-audio. Fig. 9 shows the SNR of proposed method, Chang *et al.*'s method [19] and Xiang *et al.*'s method [10] at different embedding rates. It is not difficult to find that the proposed method has a significant improvement in SNR compared with other two methods. This is because the proposed method modifies the cover slightly.

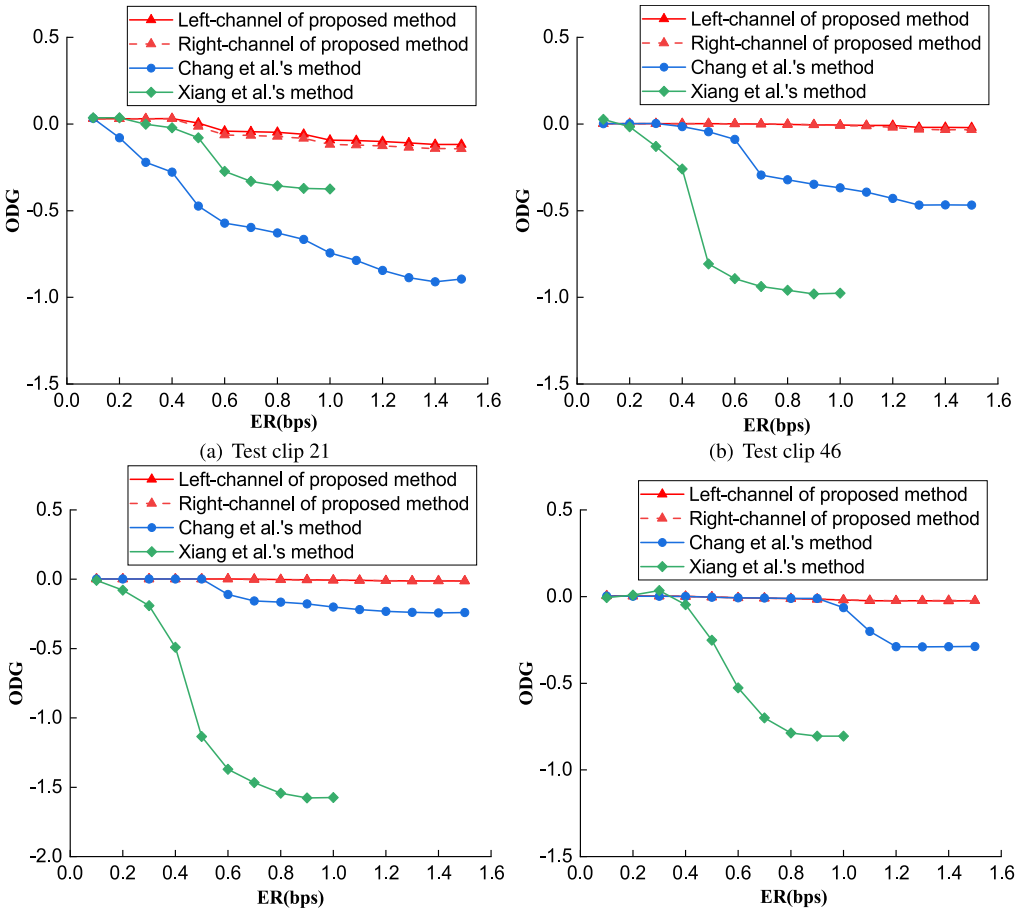


**FIGURE 10.** ODG value for proposed method, Xiang [10] and Chang [19] on single instruments (clip 21), vocal (clip 46), speech (clip 54) and orchestra (clip 66). The left channel and right channel audio generated by our method is very similar (the two red lines are almost overlapped).

In order to evaluate the auditory quality of a stego-audio more objectively, Fig. 10 shows the ODG of proposed method, Chang *et al.*'s method [19] and Xiang *et al.*'s method [10] at different embedding rates. It is observed that our method performs better than the other two methods on ODG for different audio. In addition, it can be seen that the quality of left channel audio is similar to right channel audio, which conforms to the characteristics of dual-channel audio. Finally, Table 3 and Table 4 show the SNR and ODG results of nine kinds of audio respectively. Above all, compared with other methods, the proposed approach for a variety of audio has universality, and performance are superior to other methods.

## V. CONCLUSION

This article presents a novel audio reversible data hiding scheme. The proposed method can embed three secret bits into one sample pair. Firstly, we convert binary secret data into the novenary digits. Then embedding the digits into the original single-channel audio using magic matrix and generates two single-channel stego-audio can be combined into a dual-channel audio. The human ear cannot distinguish the difference between the original single-channel audio and the dual-channel audio. In addition, compared with other digital media, the dual-channel audio cover is prevalent in cyberspace, which is more natural than two consecutive images. Experimental results show the proposed method has lower distortion and the audio quality is better than the existing algorithms under the same embedding rate.

## REFERENCES

[1] C.-K. Chan and L. M. Cheng, "Hiding data in images by simple LSB substitution," *Pattern Recognit.*, vol. 37, no. 3, pp. 469–474, Mar. 2004.
[2] A. Westfeld, "F5—A steganographic algorithm," in *Proc. Int. Workshop Inf. Hiding*. Berlin, Germany: Springer, 2001, pp. 289–302.
[3] T. Filler, J. Judas, and J. Fridrich, "Minimizing additive distortion in steganography using syndrome-trellis codes," *IEEE Trans. Inf. Forensics Security*, vol. 6, no. 3, pp. 920–935, Sep. 2011.
[4] Y.-Q. Shi, X. Li, X. Zhang, H.-T. Wu, and B. Ma, "Reversible data hiding: Advances in the past two decades," *IEEE Access*, vol. 4, pp. 3210–3237, 2016.
[5] A. Khan, A. Siddiqa, S. Munib, and S. A. Malik, "A recent survey of reversible watermarking techniques," *Inf. Sci.*, vol. 279, pp. 251–272, Sep. 2014.
[6] D. Hou, W. Zhang, J. Liu, S. Zhou, D. Chen, and N. Yu, "Emerging applications of reversible data hiding," in *Proc. 2nd Int. Conf. Image Graph. Process. (ICIGP)*, 2019, pp. 105–109.

[7] J. Tian, "Reversible data embedding using a difference expansion," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 8, pp. 890–896, Aug. 2003.
[8] F. Peng, X. Li, and B. Yang, "Adaptive reversible data hiding scheme based on integer transform," *Signal Process.*, vol. 92, no. 1, pp. 54–62, Jan. 2012.
[9] A. Nishimura, "Reversible audio data hiding using linear prediction and error expansion," in *Proc. 7th Int. Conf. Intell. Inf. Hiding Multimedia Signal Process.*, Oct. 2011, pp. 318–321.
[10] S. Xiang and Z. Li, "Reversible audio data hiding algorithm using non-causal prediction of alterable orders," *EURASIP J. Audio, Speech, Music Process.*, vol. 2017, no. 1, p. 4, Dec. 2017.
[11] X. Li, B. Li, B. Yang, and T. Zeng, "General framework to Histogram-Shifting-Based reversible data hiding," *IEEE Trans. Image Process.*, vol. 22, no. 6, pp. 2181–2191, Jun. 2013.
[12] W.-L. Tai, C.-M. Yeh, and C.-C. Chang, "Reversible data hiding based on histogram modification of pixel differences," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 19, no. 6, pp. 906–910, Jun. 2009.
[13] P. Tsai, Y.-C. Hu, and H.-L. Yeh, "Reversible image hiding scheme using predictive coding and histogram shifting," *Signal Process.*, vol. 89, no. 6, pp. 1129–1143, Jun. 2009.
[14] X. Li, J. Li, B. Li, and B. Yang, "High-fidelity reversible data hiding scheme based on pixel-value-ordering and prediction-error expansion," *Signal Process.*, vol. 93, no. 1, pp. 198–205, Jan. 2013.
[15] F. Peng, X. Li, and B. Yang, "Improved PVO-based reversible data hiding," *Digit. Signal Process.*, vol. 25, pp. 255–265, Feb. 2014.
[16] X. Wang, J. Ding, and Q. Pei, "A novel reversible image data hiding scheme based on pixel value ordering and dynamic pixel block partition," *Inf. Sci.*, vol. 310, pp. 16–35, Jul. 2015.
[17] S. Weng, Y. Shi, W. Hong, and Y. Yao, "Dynamic improved pixel value ordering reversible data hiding," *Inf. Sci.*, vol. 489, pp. 136–154, Jul. 2019.
[18] C.-C. Chang, T. Kieu, and Y.-C. Chou, "Reversible data hiding scheme using two steganographic images," in *Proc. TENCON - IEEE Region 10 Conf.*, Oct. 2007, pp. 1–4.
[19] C.-C. Chang, T.-C. Lu, G. Horng, Y.-H. Huang, and Y.-M. Hsu, "A high payload data embedding scheme using dual stego-images with reversibility," in *Proc. 9th Int. Conf. Inf., Commun. Signal Process.*, Dec. 2013, pp. 1–5.
[20] C. Qin, C.-C. Chang, and T.-J. Hsu, "Reversible data hiding scheme based on exploiting modification direction with two steganographic images," *Multimedia Tools Appl.*, vol. 74, no. 15, pp. 5861–5872, Aug. 2015.
[21] L.-P. Chi, C.-H. Wu, and H.-P. Chang, "Reversible data hiding in dual stegano-image using an improved center folding strategy," *Multimedia Tools Appl.*, vol. 77, no. 7, pp. 8785–8803, Apr. 2018.
[22] G.-D. Su, Y. Liu, and C.-C. Chang, "A square lattice oriented reversible information hiding scheme with reversibility and adaptivity for dual images," *J. Vis. Commun. Image Represent.*, vol. 64, Oct. 2019, Art. no. 102618.
[23] X. Huang, N. Ono, I. Echizen, and A. Nishimura, "Reversible audio information hiding based on integer DCT coefficients with adaptive hiding locations," in *Proc. Int. Workshop Digit. Watermarking*, 2013, pp. 376–389.
[24] E. Committee. (2008). *Sound Quality Assessment Material Recordings for Subjective Tests*. [Online]. Available: https://tech.ebu.ch/publications/sqamcd

**HENG YU** received the B.Eng. degree from the Applied Science College, Jiangxi University of Science and Technology, in 2018. He is currently pursuing the master's degree with the Faculty of Electrical Engineering and Computer Science, Ningbo University. His research interests include data hiding and information security.

**RANGDING WANG** (Member, IEEE) received the Ph.D. degree from Tongji University, in 2004. He is currently a Full Professor with Ningbo University, China. His research interests include multimedia security, digital watermarking for digital rights management, data hiding, and steganography.

**LI DONG** (Member, IEEE) received the B.Eng. degree from Chongqing University, in 2012, and the M.S. and Ph.D. degrees from the University of Macau, in 2014 and 2018, respectively. He is currently an Assistant Professor with the Department of Computer Science, Faculty of Electrical Engineering and Computer Science, Ningbo University. His research interests include statistical image modeling and processing, multimedia security and forensic, and machine learning.

**DIQUN YAN** (Member, IEEE) received the B.S., M.S., and Ph.D. degrees from Ningbo University, in 2002, 2008, and 2012, respectively. He was a Visiting Scholar with the New Jersey Institute of Technology, Newark, NJ, USA, from 2014 to 2015. He is currently an Associate Professor with the Faculty of Electrical Engineering and Computer Science, Ningbo University, China. He is also the Head of the Computer Science Department. His current research interests include speech processing and multimedia forensics.

**YONGKANG GONG** received the B.Eng. degree from Ningbo University, in 2018. He is currently pursuing the master's degree with the Faculty of Electrical Engineering and Computer Science, Ningbo University. His research interests include adversarial machine learning and information security.

**YUZHEN LIN** received the B.Eng. degree from China Jiliang University, in 2017. He is currently pursuing the master's degree with Ningbo University. His current research interests include steganalysis and steganography.