



Improving the Performance of Data Hiding by Designing Three Samples-based Smoothing in Audio

Ryan Setiawan^{1*} Tohari Ahmad¹

¹*Department of Informatics, Institut Teknologi Sepuluh Nopember, Indonesia*

* Corresponding author's Email: ryan.stwn93@gmail.com

Abstract: The internet is widely applied to facilitate the transferring process of data as a means of communication. Some of this transferred message has confidential information and must be protected. One of the techniques to protect is data hiding. It uses carrier media as a cover to hide information inside it, one of which is audio. However, there are some problems in the use of audio, namely the limited capacity and sound quality after data insertion. In this paper, we address these problems by interpolating the audio samples using Modification Lagrange Interpolation (MLI) and exploring Reduce Difference Expansion (RDE) for embedded the secret. Next, samples are smoothed to improve the quality based on three sampling signal points of the stego audio. The testing of proposed method is carried out on fifteen mono audio data sets from various genres the covers and ten files as confidential information data with various file sizes. The results show that using the proposed methods, the performance increases that the average quality of each audio genre raises by about 4dB and able to increase the capacity of message data that can be hidden by 20%.

Keywords: Steganography, Interpolation, RDE, Smoothing signal, Data protection, Data hiding.

1. Introduction

Sending information over the internet with one another is very easy using available services provided today. Much information may be hidden using any data hiding technique to guarantee its security. For this purpose, various cover media can be used to carry the information, such as images, audio, video, and text documents. Some data hiding techniques are reversible, for example steganography, cryptography, and watermarking. Steganography is an art and science of calculation to hide data information and maintain the information that has been well protected. The cover media itself should be similar between before and after the insertion process.

There are some opportunities for improvements to get better quality and larger capacity for the carrier media [1], especially in audio media by utilizing signals that can be manipulated. Thus, as in other research, data insertion can be done.

Research about audio covers has been developed, which some of them only focus on one challenge, either quality or capacity. To improve the quality,

there is some method to modify the bits of the audio cover layer by Least Significant Bit (LSB) [2], while applies the Reversible Data Hiding (RDE) scheme by inserting histogram shifting data of the encrypted cover media domain [3]. They implement a method, which partitions the cover into mutation blocks of 2×2 and 3×3 without data expansion. The experimental results show that the encrypted cover media and histogram shifting algorithm is able to achieve high enough embedding capacity, compared with the similar algorithm. Other research introduces the insertion of information using a modification of the LSB to change the maximum audio signal from randomly generated data. This is also gives more spaces to achieve higher capacity [4]. Andra et al. [5] combines Reduced Difference Expansion (RDE) method with Generalized Difference Expansion to increase the capacity of the audio cover media when inserting the data. Some research can improve the performance significantly, for example that of [6], which is able to guarantee an increase of both capacity and quality using the signal interpolation method and LSB based on the difference between the

original signal and the new signal from Newton Divided Difference Interpolation (NDDI). Furthermore, this research can still be continued to maximize the capacity when a message was inserted and still maintain good quality of the audio cover.

In this paper we propose a method for inserting confidential data using the MLI technique that is applied to a single-channel audio cover by exploring the RDE hiding data insertion method. It is to increase the capacity and designed a smoothing algorithm model based on three sampling point to maintain the quality of the stego audio. The results are then passed through signal interpolation stage and file insertion performed on the audio sample. Thus, the audio only adjusted the sample rate not to change the duration of the audio playback and sound looks similar as the audio cover before the data insertion was performed.

This paper is organized as follows. Section 2 describes the previous research which relates to our work. The proposed method is explained in Section 3, and followed by the experimental result in Section 4. Lastly, the conclusion of this paper is drawn in Section 5.

2. Related works

There are several previous research related to audio interpolation sampling methods, data insertion methods, and smoothing process. To reduce distortion in stego audio at the sampling rate with insertion and extraction of digital data in the audio cover using LSB embedding techniques based on the results of the Human Auditory System (HAS) [7]. The steganography process must eliminate the effects of suspicion of stego audio results. This research [7] uses the Nyquist-Shannon sampling theorem to obtain maximum sampling results, but this embedding method for the capacity is relative low.

Steganography as discussed in this paper is inseparable from the importance of location maps as the container selection process, so that the messages can be inserted. There are two ways to carry the location map information: to send it separately from the cover and the information is inserted into the cover itself. Research by Al-Hooti and Ahmad [8] proposed inserting data using a new frequency transformation method with audio based on Discrete Cosine Transform (DCT) and location maps. It helped to fully restore the secret message and cover from the frequency sampling that increases or decreases by 10 to -10. The results can guarantee 100% recovery of secret messages and audio cover, while the SNR is about 75dB. However, this mechanisms is needed to send two files the stego

audio, and the location maps. Then, Assini et al. [9] combined Discrete Wavelet Transform (DWT), DCT, and Singular Value Decomposition (SVD). This technique is more complex, through a watermark insertion approach to cover media images so they are not visible. This data insertion process is applied to a single value from the watermark that is entered into the singular value of the highest frequency sub-band of the cover. The simulation shows that this hybrid method is able to provide sturdy values for attacks such as Gaussian noise, Wiener filters and others.

Research of Jung and Yoo [10] also used images, and proposed interpolation scale values from neighbours for each pixel. A simple algorithm is used to find the source of adjacent pixels to produce new pixel values, where the data insertion sequence can run zig-zag, left-to-right, even top-down by inserting pixel data using $\log_2|d|$. The results show an increase of 35dB compared to reversible data hiding methods.

The research is continued by Jan et al. [11] by interpolating the cover images, and divided into 2×2 to prevent overlapping blocks. Next, it works on the capacity of the pixel n_i , where $n_i = \lfloor \log_2|d_i| \rfloor$, $i = 1,2,3$. The comparison results show that the capacity value are the same. However, the quality of the cover is much better. Development is continued by Jung and Yoo [12] in order to increase the capacity and quality of cover images, by using LSB. It sets $k = 2$ and $k = 3$, where k is the rightmost least significant bits value to replace the zigzag interpolated pixel values resulted by each block. That research uses $k = 2$. This depicted that the research is superior in terms of the quality, and $k = 3$ delivers the greater maximum capacity. Although this method has a better quality and capacity, the number of specified bits has become its weakness, which has an effect on the overall performance.

Ahmad and Fiqar [13] proposed to use NDDI method as a substitute for locations maps from the curve forming process of a cover file. The interpolation method also acts as a pivoting point for the interpolated sampling signal to be inserted by using allocation numbers are created from the different between two samples and deigning a new smoothing signal to maintain the quality of stego audio. This is a main reference to improve the capacity by using some various high-order interpolation; and to refine the quality by improving the smoothing signal of the previous method [13]. For this purpose, we design it based on three-point sampling signals after the embedding.

In this paper we use the RDE method to insert data, which adopts the value of the neighbouring signal sampling. Some research has used this RDE insertion

method, such as the research of Yi et al. [14]. It can improve a method called IRDE by exploring reversible data hiding for digital images using location maps to reduce interference or noise when the file is extracted. The embedded data can be extracted correctly and the original image cover and information message is recovered without distortion.

Other research, which is proposed by Kurniawan et al. [15], combines pre-processing of modulus 3 for secret messages, so that they can be prepared into the image cover. An RDE-based method from Liu et al. [16] is influenced by the neighbouring samples to determine the results of the reduction, which the data are inserted into the new signals after reduction. Therefore, here is h in Eq. (1), to determine the different samples neighbour.

$$h = x - y \tag{1}$$

To reduction ratio of the neighbour sampling values using Eq. (2), (3), (4), with notation x and y .

$$h' = h - 2^{\lfloor \log_2 h \rfloor} \tag{2}$$

$$h'' = 2 \times h' + b \tag{3}$$

$$l = \left\lfloor \frac{x+y}{2} \right\rfloor \tag{4}$$

Notation h' is the value of result that divided to proceed with Eq. (2), which h' is multiplied and added to the calculated bit value in Eq. (3), From these results it can then be continued with Eq. (4), notation l is the new value of reduce different that will be used for the RDE insertion method with 3 condition trends.

3. Proposed method

Generally, there are two essential parts in the proposed method: encoding and decoding processes, or also called embedding or inserting and extraction, respectively.

The design of payload insertion is done at a new sample value that is inserted continuously using the RDE method from the Polynomial Lagrange Interpolation signal location. Based on [17] and [18] that use 4 bits of Most Significant Bits (MSB), the payload is to be inserted. So, we get Eq. (5), to divide the length of the secret message value to be inserted.

$$b = 4bit\ MSB \tag{5}$$

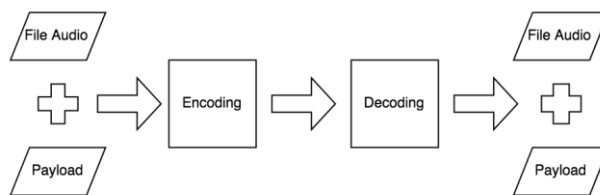


Figure. 1 General concept of data hiding

A. Encoding process

The encoding process is explained in Fig. 2, where the 16-bit audio cover sampling resulted from the interpolation value is a substitute for the location map. This process generates samples $2n - 1$ as a new insertion space. As in other interpolation-based method, it can increase the amount of samples. This method is designed for *.wav audio cover because this audio type does not suffer from losing information.

The initial step needs to be done by the normalization. It is to change the value of the sampling audio cover file data type from an integer value (from -32768 to 32767) to an unsigned integer (from 0 to 65535), so all audio signal magnitudes are positive. An example of this process is illustrated as follows. It is assumed we have 5 signal samplings; those are $y_1 = 80, y_2 = 84, y_3 = 98, y_4 = 94, y_5 = 90$ be the value of the original audio samples.

Then, the MLI implementation is done by using Eq. (6), (7). New MLI signals are used as pivoting points or location maps to insert the payload data. So, the original data are not interfere with the interpolation audio signal. At this stage, the original audio signal are increase to the MLI results. We find that the y_1 sample does not change, y_2 has become y_3 , and y_4 become y_5 so on. Then the audio signal is in odd and even index.

$$f_1(y) = \frac{(y - y_1)}{(y_0 - y_1)}f(y_0) + \frac{(y - y_0)}{(y_1 - y_0)}f(y_1) \tag{6}$$

$$f_2(y) = \frac{(y - y_1)(y - y_2)}{(y_0 - y_1)(y_0 - y_2)}f(y_4) + \frac{(y - y_0)(y - y_2)}{(y_1 - y_0)(y_1 - y_2)}f(y_5) + \frac{(y - y_0)(y - y_1)}{(y_2 - y_0)(y_2 - y_1)}f(y_7) \tag{7}$$

From Eqs. (6) and (7), the notation $f_n(y)$ is the new MLI signal, and here is an example can be calculated for explaining interpolation signal.

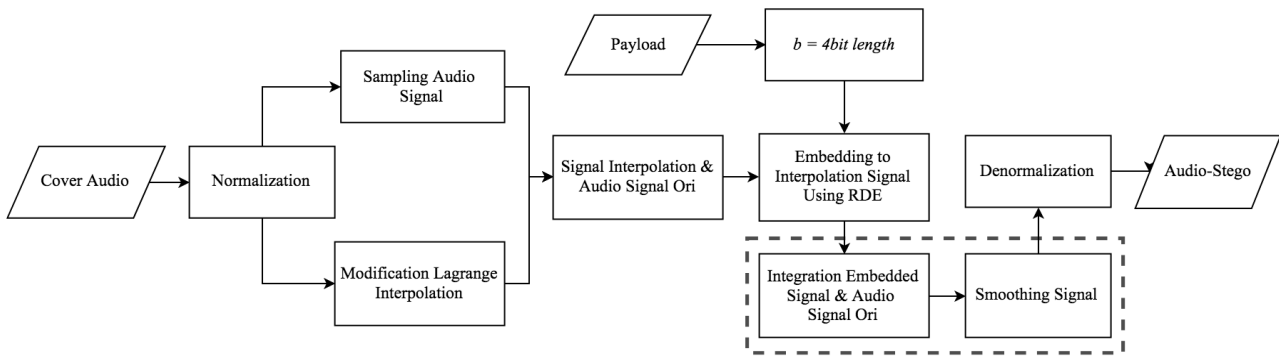


Figure. 2 The embedding process

$$f_1(2) = \frac{(2-3)}{(1-3)}80 + \frac{(2-1)}{(3-1)}84$$

$$f_1(2) = 82$$

$$f_2(6) = \frac{(6-5)(6-7)}{(4-5)(4-7)}98 + \frac{(6-4)(6-7)}{(5-4)(5-7)}94$$

$$+ \frac{(6-3)(6-5)}{(7-3)(7-5)}90$$

$$f_2(6) = 91.1$$

$$h'_1 = 2 - 2^{\log_2 2} = 0$$

$$h'_2 = 2.4 - 2^{\log_2 2.4} = 0$$

$$h''_1 = 2 \times 0 + 2 = 2$$

$$h''_2 = 2 \times 0 + 7 = 7$$

$$l_1 = \frac{82 + 80}{2} = 81$$

$$l_2 = \frac{91.6 + 94}{2} = 92.8$$

The integrity of the MLI signal and the original audio signal in this process are combined to get a new vector value of the combined as in Eq. (8).

$$f_n(y) = [y_1, \bar{y}_2, y_3, y_4, y_5, \bar{y}_6, y_7, \dots, y_n, (y_{n+1})] \quad (8)$$

Notation y_n is the original audio signal while \bar{y}_n is an MLI signal. At this stage, a new signal is obtained [83, 93]. The input of the payload is in the form of text, for example, from the payload "sec" we get "s = (00100111)".

$$b_n = \text{bit2dec}(4\text{bits})$$

Based on this bit string, we have:

$$b_1 = 0010$$

$$b_2 = 0111$$

These blocks of bits are then converted to decimal value, $b_1 = 2$, and $b_2 = 7$. The process of embedding to MLI signals uses the RDE method. There are three conditions of the signal trend resulted from the audio interpolation signal and its neighbours. For the first we looking for difference values.

$$h_1 = |82 - 80| = 2$$

$$h_2 = |91.6 - 94| = 2.4$$

If the sampling value y_n is smaller than y_{n+1} , then the first condition is used. However, if sampling value y_n is greater than y_{n+1} the second condition is used. In case those values are the same, the third condition is applied. The conditions are explained in Eq. (9), for each data insertion.

$$\bar{y}'_n = \begin{cases} l + \left\lfloor \frac{h''}{2} \right\rfloor + 1, & \text{if } b \neq 0 \text{ and } y_n < y_{n+1} \\ l - \left\lfloor \frac{h''}{2} \right\rfloor, & \text{if } b \neq 0 \text{ and } y_n > y_{n+1} \\ l + \left\lfloor \frac{h'' + b}{2} \right\rfloor, & \text{if } b = 0 \text{ and } y_n = y_{n+1} \end{cases} \quad (9)$$

According to Eq. (9), the high magnitude of the audio signal is used as a reference for inserting the MLI signal. The embedding method RDE is able to insert a payload to create a new value that is relatively not too much different from the neighbour value. So the embedded signal integration is generated by Eq. (8). The signal input is [80, 82, 84, 98, 94, 92, 90], for the example, we explain that the original signal that has been inserted to new the signal embedding [80, **83**, 84, 98, 94, **89.3**, 90], it is obtain.

$$\bar{y}'_2 = 81 + \frac{2}{2} + 1 = 83$$

$$\overline{y_6'} = 92.8 - \frac{7}{2} = 89.3$$

Furthermore, the stego signal is smoothed to get a lower difference after the RDE embedding process. In this method, the neighbouring signals are to be the reference for this smoothing step, which is provided in Eq. (11). It uses 3 adjacent samples resulting from the development of Eq. (10), whose details are explained in Fig. 5.

$$Y_n = \frac{Y_{n-1} + \overline{y_n'} + Y_{n+1}}{3} \tag{10}$$

$$Y_{j,n} = \frac{Y_{n-1} + (3^j - 2) \times \overline{y_n'} + Y_{n+1}}{3^j} \tag{11}$$

Index (j) is the reduction level in Eq. (11). In this case, higher value of j means that more smoothing to the embedded signal is applied. That is, $j = 1, 2, 3, \dots$ and so on produces 50%, 75%, 87%, respectively, of reduce the difference. Index (n) is the sample of the point to be reduced, and index $Y_{j,n}$ is a new reduced signal which is influenced by the neighbour signal. In this case, Y_{n-1} and Y_{n+1} are determined as a neighbour signal from the sampling MLI. For example, by using the index level (j) = 1, the signal quality reduces by 50%. The greater the index value (j) the greater the value of the divisor. The resulted signal can be depicted as follows. [80, **83**, 84, 94, **89.3**, 90] → 80, **82.3**, 84, 94, **91.1**, 90]. By using Eq. (11), we get following values.

$$\begin{aligned} \overline{y_1'} = 83 \rightarrow Y_{1,1} &= \frac{80 + (3^1 - 2) \times 83 + 84}{3^1} \\ &\rightarrow y_{1,1} = 82.3 \end{aligned}$$

$$\begin{aligned} \overline{y_2'} = 93 \rightarrow Y_{1,1} &= \frac{94 + (3^1 - 2) \times 89.3 + 90}{3^1} \\ &\rightarrow y_{1,2} = 91.1 \end{aligned}$$

So, the reduction results will make the signal value closer to its neighbours, which makes the quality of the signal improves. Furthermore, the denormalization process is to returns the position of the magnitude value from unsigned integer to integer value. This process is needed by the encoder to make it readable. The sampling rate almost double from the number of the original audio samples, which is $2 \times (n - 1)$. Furthermore, the initial sampling rate of 44100 Hz changes to 88200 Hz. At this stage, stego audio can be played on the audio player.

B. Decoding process

Decoding process is to extract the payload that has been inserted, as provided in Fig. 4. The first step is normalization to make the split signals readable to the audio encoder. Meanwhile, the sampling rate is changed back to the initial rate, which is about half of the existing. After that, the embedded we used here to continue pre-smoothing in this extraction process the smoothing signal affected by neighbouring signal values as well, so that no information is left behind. Then this pre-smoothing process uses the derivative by Eq. (12), we obtain.

$$\overline{y_{j,n}''} = 3 \times Y_{j,n} - (y_{n-1} + y_{n+1}) \tag{12}$$

The notation $\overline{y_n''}$ in Eq. (12) is a new sample signal after extraction, and index (j) in pre-smoothing must be the same as the encoding process. By following the previous example, we have input the smoothing signal [80, **82.3**, 84, 94, ,98, **91.1**, 90], then by employing Eq. (11), in this case smoothing level $j = 1$. Here, we find by Eq. (12), produces a signal value [80, 83, 84, 94, 98, 89.3, 90].

$$\overline{y_{1,2}''} = 3 \times 82.33 - (80 + 84) = 83$$

$$\overline{y_{1,6}''} = 3 \times 91.6 - (94 + 90) = 89.3$$

Then, according to embedded sampling extraction is continued using the RDE method by Eq. (9). By following those requirements, this extraction leads to.

$$b_1 = 83 - 81 = 2$$

$$b_2 = (92.8 - 89.3) \times 2 = 7$$

At this step the payload is successfully recovered, which is $(2)_{10}$ and $(7)_{10}$. In accordance with Eq. (5), if the decimal is converted to binary then the payload information $b_1 = (0010)_2$ and $b_2 = (0111)_2$. The payload are combined into the bit stream to be reconstructed, so "s = $(00100111)_2$ " is the information inserted in the stego audio and the process ends when the entire payload is successfully recovered.

4. Experiment result and analysis

Data set we used for the experiment is from Instrument Recognition in Musical Audio Signals (IRMAS) [19]. It is public and widely used in various research for free.

There are 15 audio files to use in this research, comprising 5 different instruments and are divided into 3 genres. As explained in detail in Table 1, the

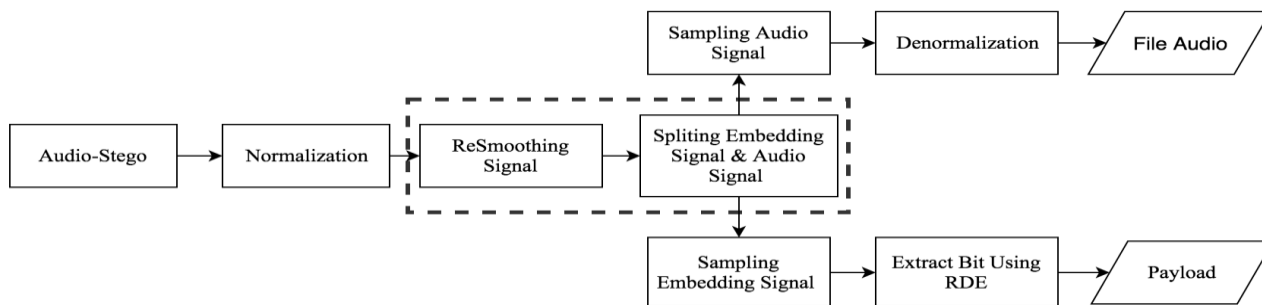


Figure. 3 The extraction process

Table 1. Audio dataset experiment

Name	Genre	Instrument
Audio1	Country-Folk (cou_fol)	Cello (cel)
Audio2	Classical (cla)	Cello (cel)
Audio3	Pop-Rock (pop-roc)	Cello (cel)
Audio4	Country-Folk (cou_fol)	Acoustic guitar (gac)
Audio5	Classical (cla)	Acoustic guitar (gac)
Audio6	Pop-Rock (pop-roc)	Acoustic guitar (gac)
Audio7	Country-Folk (cou_fol)	Piano (pia)
Audio8	Classical (cla)	Piano (pia)
Audio9	Pop-Rock (pop-roc)	Piano (pia)
Audio10	Country-Folk (cou_fol)	Saxophone (sax)
Audio11	Classical (cla)	Saxophone (sax)
Audio12	Pop-Rock (pop-roc)	Saxophone (sax)
Audio13	Country-Folk (cou_fol)	Voice (voc)
Audio14	Classical (cla)	Voice (voc)
Audio15	Pop-Rock (pop-roc)	Voice (voc)

Table 2. Characteristic cover audio

No	Information	Character
1	Media cover	Audio 1 ~ 15
2	Channels	Single - Layer
3	Sample Rate	44100
4	Sampling Rate Stego	88200
5	Duration	~3 Second
6	Sample Bit	16 Bit

characteristics of audio is in Table 2, whereas the payload file itself consists of 10 files with different

sizes, ranging from 10kb, 15kb, 20kb, 25kb, 30kb, 35kb, 40kb, 46kb, 51kb and 56kb.

Then evaluate the measurement comparison between 2 audio cover media, namely the file that has been inserted with the interpolated cover media. The Measurements were made using the Signal to Noise Ratio (SNR) described in Eq. (13), to determine the quality value.

$$SNR(s_o, s_w) = 10 \log_{10} \frac{\sum_n [s_o(n)]^2}{\sum_n [s_w(n) - s_o(n)]^2} \quad (13)$$

Notation s_o is the original signal from the audio cover while the signal is from the insertion of payload data, and $\sum_n [s_o(n)]^2$ is a total of the total value of the total original data n as squared divided by $\sum_n [s_w(n) - s_o(n)]^2$. The difference between each audio cover data that has been inserted and the cover of the stego audio squared results.

So, in this paper there are two parts of the first method evaluation of determining the maximum value of the insertion capacity of audio cover. The results of audio quality through calculation SNR described in dB. The data are tested using several types of audio genres as a comparison resulting from variations the amount of data payload. Thus, affecting the value of the results of the quality of stego audio before and after payload has been inserted.

The capacity of the Audio1-Audio15 cover is explained in Fig. 4, while y -axis represent the number of bits that can be inserted and the SNR of each audio cover then x -axis represents audio cover data. This insertion is influenced by the availability of space from the MLI signal, the number of sampling signals after interpolation is increasing by $2n - 1$ and this affects the amount the available of room space for the payload data to be inserted. The comparison is done by following the scenario of [12] and [13] to compare with the proposed method through 4 times testing, using low-order and high-order interpolation.

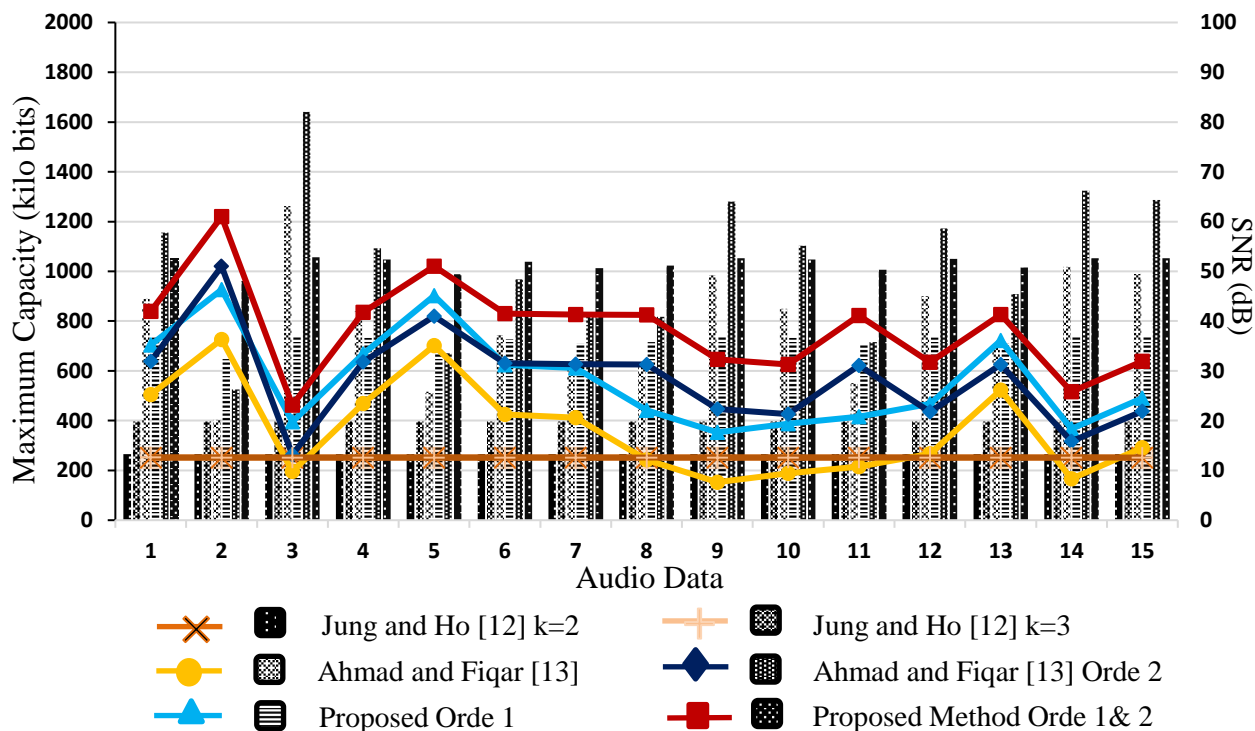


Figure. 4 Comparison of the embedding capacity between Jung and Yoo [12] k=2 and k=3, Ahmad and Fiqar [13] orde 1, Proposed method Orde 1, Ahmad and Fiqar [13] Orde 2, Proposed Method using MLI signal (Orde 1 & 2)

This evaluation shows that the overall use of high-order MLI signals can increase the value of the capacity based on the previous methods [12] and [13].

For the comparison, the capacity increases around 246% from [12], and around 20% from [13]. This increase of the capacity is achieved by either lengthening the audio sampling rate which leads to increasing the number of samples. However, when the experiment was carried out using the same orde, the proposed capacity value decreased by 10% for low-order interpolation and 4% for high-order interpolation. It is because the RDE method has limited insertion terms to maintain the stability of the quality value of the audio cover and the SNR value shows the opposite of the increase the maximum capacity for overall audio. Basically, this capacity value follows the instrument and genre of all audio files.

In this proposed method, the increase value can be found in the genre Classical music provided in Fig. 4, especially in Audio2 and Audio14. It is evident in terms of the capacity that shows the higher and the lowest graph than the comparison method. However, the capacity value actually has the same pattern with [13] that is follows the instrument and genre audio cover.

In the experiment to measure the quality of the audio data after the smoothing process, we analysed the results shown in Fig. 5. Those data are the average of the SNR value after the audio cover has been inserted by various payload sizes. Similar to that of the capacity, the measurement of the quality is done by following the scenarios in [13]. It is to compare [12] and [13] with the proposed method through 4 times testing using low-order and high-order interpolation.

The evaluation shows that RDE and smoothing methods can improve the quality about 4dB for all testing interpolation in [13]; and it can improve about 6dB for the value of k . The result showed that the the best quality of the proposed method is achieved by the classical genre. This is because the classical has relatively a low difference of sample signals. Meanwhile, the lowest quality value is in cello and pop-rock genres, which means that the difference between successive sample is relatively high. However, the increase in SNR value is inversely proportional to the insertion capacity for all audio covers. So, the sampling method in high order MLI able to increase the capacity of the signal magnitude generated by Eq. (6), (7). Then, RDE is able to insert the payload by three conditions specified by the

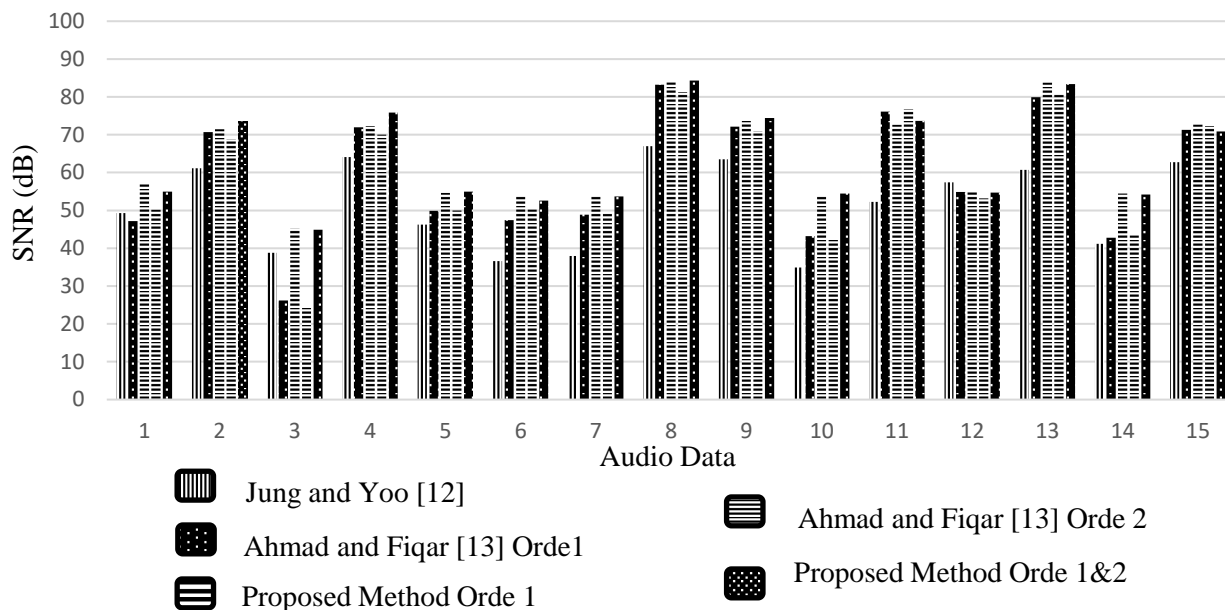


Figure. 5 Comparison of the quality between Jung and Yoo [12], Ahmad Fiqar [13] orde 1, Proposed method Orde 1, Ahmad and Fiqar [13] Orde 2, Proposed Method Orde 1 & 2

trends described in Eq. (9). And finally, the proposed smoothing method signal is able to reduce the difference between the embedded signal and its neighbours, y_{n-1} and y_{n+1} . This has been able to make it more similar to them.

5. Conclusion

According to the experimental results in this paper, we can conclude that audio sampling using the MLI method has been successfully carried out and is able to generate an increase of the capacity. Then, the described insertion method is able to maximize the number of confidential data to be put into the audio and the designed smoothing method can maintain the quality of the stego audio.

Furthermore, the high-order interpolation is also able to raise the capacity quite well in accordance with the addition of the new sampling signal. In this research, the MLI is implemented with order 1 & 2, which means the space for the embedding has been gone up.

It is also depicted that the RDE can fully embed the secret into the interpolated signals. The possible noises are minimized by the proposed smoothing algorithm. It has been proven in the experiment that this scheme can maintain the quality of the generated stego data.

For the evaluation purpose, 10 variations of payload data have been inserted into 15 audio covers. The highest SNR value seen in Audio8 the Country-Folk music with genre human singing voice, and for

the capacity value Audio3 Pop-Rock with genre Cello instruments has a maximum amount room space of 1057344 bits.

There are some rooms to improve the capacity of this method. It is done by combining with other embedding methods to further expand the size of data to hide. Besides, it is possible to extend the number of bits that can be determined in the value of the sampling signal.

Acknowledgments

This research is supported by the postgraduate program of Institut Teknologi Sepuluh Nopember, Indonesia.

References

- [1] I. Jawad, P. Premaratne, P. James, and B. Halloran, "Comprehensive Survey of Image steganography: Techniques, Evaluations, and Trends in Future Research", *Neurocomputing*, Vol. 335, pp. 299–326, 2019.
- [2] B. Datta, P. K. Pal, and S. K. Bandyopadhyay, "Multi-Bit Data Hiding in Randomly Chosen LSB Layers of an Audio", In: *Proc. of 2016 15th International Conf. Information Technology ICIT 2016*, pp. 283–287, 2017.
- [3] R. Zhang, C. Lu, and J. Liu, "A High Capacity Reversible Data Hiding Scheme for Encrypted Covers based on Histogram Shifting", *Journal Information Secure Application*, Vol. 47, pp. 199–207, 2019.

- [4] S. N. Choudhury, "Enhanced Data Hiding Audio Steganography Using Signal Ramping", In: *Proc. of 2018 Fourth International Conf. Research Computational Intelligence Communication Networks*, pp. 258–263, 2018.
- [5] M. B. Andra, T. Ahmad, and T. Usagawa, "Medical Record Protection with Improved Grde Data Hiding Method on Audio Files", *English Letters*, Vol. 25, No. 2, pp. 112–124, 2017.
- [6] J. Liu, K. Zhou, and H. Tian, "Least-Significant-Digit Steganography in Low Bitrate Speech", In: *Proc. of IEEE International Conf. Communication and Information System Security Symposium*, pp. 1133–1137, 2012.
- [7] R. R. Devi and D. Pugazhenth, "Ideal Sampling Rate to Reduce Distortion in Audio Steganography", *International Conf. on Computational Modeling and Security*, Vol. 85, pp. 418–424, 2016.
- [8] M. H. A. Al-Hooti and T. Ahmad, "Reversible Data Hiding in Audio Based on Discrete Cosine Transform and Location Maps", *International Journal of Intelligent Engineering and Systems*, Vol. 12, No. 3, pp. 41–49, 2019.
- [9] I. Assini, A. Badri, K. Safi, A. Sahel, and A. Baghdad, "A Robust Hybrid Watermarking Technique for Securing Medical Image", *International Journal of Intelligent Engineering and Systems*, Vol. 11, No. 3, pp. 169–176, 2018.
- [10] K. H. Jung and K. Y. Yoo, "Data Hiding Method Using Image Interpolation", *Computer Standarts and Interfaces*, Vol. 31, No. 2, pp. 465–470, 2009.
- [11] S. Jan, S. J. Hsu, and C. Chiu, "An Improved Data Hiding Method Using Image Interpolation", In: *Proc. of 2011 Seventh International Conf. Intellignet Information Hiding and Multimedia Signal Processing*, No. 1, pp. 185–188, 2011.
- [12] K. H. Jung and K. Y. Yoo, "Steganographic Method based on Interpolation and LSB Substitution of Digital Images", *Multimedia Tools Applications*, Vol. 74, No. 6, pp. 2143–2155, 2015.
- [13] T. Ahmad and T. P. Fiqar, "Enhancing The Performance of Audio Data Hiding Method by Smoothing Interpolated Samples", *International Journal Innovative Computing Information and Control*, Vol. 14, No. 3, pp. 767–779, 2018.
- [14] H. Yi, S. Wei, and H. Jianjun, "Improved Reduced Difference Expansion based Reversible Data Hiding Scheme for Digital Images", In: *Proc. of the 9th International Conf. on Electronic Measurement and Instruments*, pp. 4315–4318, 2009.
- [15] Y. Kurniawan, L. A. Rahmania, T. Ahmad, W. Wibisono, and R. M. Ijtihadie, "Hiding Secret Data by Using Modulo Function in Quad Difference Expansion", *International Conf. Advanced Computer Science Information Systems*, pp. 433–438, 2017.
- [16] C. L. Liu, D. C. Lou, and C. C. Lee, "Reversible Data Embedding Using Reduced Difference Expansion", In: *Proc. of the 3rd International Conf. Intelligent Information Hiding and Multimedia Signal Processing*, Vol. 1, pp. 433–436, 2007.
- [17] S. P. Rajput, K. P. Adhiya, and G. K. Patnaik, "Hide Text in Audio", In: *Proc. of 2017 International Conf. Computing, Communications, Control and Automation*, pp. 1–6, 2017.
- [18] M. Tayel, A. Gamal, and H. Shawky, "A Proposed Implementation Method of an Audio Steganography Technique", In: *Proc. of the 18th IEEE International Conf. Advanced Commucation Technology*, No. 3, pp. 180–184, 2016.
- [19] IRMAS: A Dataset for Instrument Recognition in Musical Audio Signals, <https://www.upf.edu/web/mtg/irmas/>, [Access: 3-December-2019].