

GENERATING AUDIO SIGNAL AUTHENTICATION THROUGH SECRET EMBEDDED SELF HARMONIC COMPONENT (GASASH)

Uttam Kr. Mondal¹ and J.K.Mandal²

¹ Dept. of CSE & IT,
College of Engg. & Management, Kolaghat ,
Midnapur (W.B), India
uttam_ku_82@yahoo.co.in

² Dept. of CSE,
University of Kalyani,
Nadia (W.B), India
jkm.cse@gmail.com

ABSTRACT

In this paper, an algorithm has been proposed to provide authentication of audio signal through concealment of a selective segment of its harmonic components and hiding the selective harmonic components in such a manner that will not affect over its audible quality. The replica of inner harmonic components will use for hidden signature in wavelet domain. Sampling the hidden harmonic components with the help of amplitude coding for generating lower magnitude values is the first phase of proposed technique followed by fabrication of authenticating code by embedding the lower magnitude values with selected coefficients of song signal generated via wavelet transform [symmetrization mode]. The embedded hidden self component of song signal is used to detect and identify the original song from similar available songs with the help of comparing its hidden code and respective similar inner part. A comparative study has been made with similar existing techniques and experimental results are also supported with mathematical formula based on Microsoft WAVE (".wav") stereo sound file.

KEYWORDS

Audio Signal Authentication, Average Absolute Difference (AD), Harmonic Component, Maximum Difference (MD), Mean Square Error (MSE), Normalized Average Absolute Difference (NAD), Normalized Mean Square Error (NMSE), Wavelet Transform.

1. INTRODUCTION

Enhancing audio signal security is one of the important concerns in music industries [4]. Producing duplicate version of similar audio songs with minimum expenses can choke the revenues of original investor. There for creating good quality songs with security criteria can protect these original versions and help to detect the duplicate versions. IPR of audio songs also may be protected, if concern person able to embed/ pass some secure hidden code with original versions. Therefore, combining security code without changing its audible quality level can

protect its copy right properties of original version [1, 2]. The purpose of this paper is to review this authentication problem and propose a technique to solve it.

In this paper, a framework for protecting originality of a particular song with the help of secret harmonic components with coefficient values of wavelet transform without changing its audible quality has been presented. Sampling the hidden harmonic components with the help of amplitude coding for generating lower magnitude distribution is the first phase of proposed technique followed by fabrication of authenticating code by embedding lower magnitude values with selected coefficients of song signal generated via wavelet transform [symmetrization mode]. The embedded hidden secure harmonic component as well as authenticating code will use to detect and identify the original song from similar available songs. It is experimentally observed that added extra values will not affect the song quality but provide a level of security to protect the piracy of song signal.

Organization of the paper is as follows. Generating secret code is described in the section 2.1. Sampling selected secret values set and converting into lower magnitude values (in the range of 0 to 0.1001) are presented in section 2.2. Embedding secret code is performed in section 2.3. The extraction is shown in section 2.3. Experimental results are given in section 3. Conclusions are drawn in section 4. References are given at end.

2. THE TECHNIQUE

The scheme fabricates the secure hidden code with help of wavelet transform followed by generating authenticating code. Sampling generated frequency band and Embedding Secret code (GASASH-ESC) are given in section 2.1 and section 2.2 respectively.

2.1. Generating Secret Code

Selected harmonic component of original song will use to create secret code and the generated code will be embedded with selected region of particular coefficients in wavelet domain without changing its audible quality as described in the section of 2.3. Generating secret code and selecting harmonic components are depicted in the following steps.

Step 1: Apply FFT over selected region of original song signal $f(x)$ to generate the frequency component of that region.

Step 2: Select a set of harmonic components from the output of FFT as secret key that will embed with the coefficient components in wavelet transform.

Therefore, the selected harmonic band is the secret key and converting the harmonic values into lower magnitude values as done in section 2.2 which will use for embedding with coefficients in wavelet domain as describe in section 2.3.

2.2. Sampling Secret Code

Sampling of secret code [generated by the section 2.1] is to produce set of lower magnitude values to add with the selected coefficient values of song signal in wavelet domain has been fabricated with authenticated code. Let, V is a harmonic component value [generated by applying the process as described in the section 2.1]. The magnitude of V can be represented into a set of small amplitude values as described in the following steps.

Step 1: Find individual digits of V value and represent the individual digit into binary form. After representing corresponding binary number, convert them into lower magnitude values by as described in the following example. If the magnitude value of V is 1.0051, then corresponding lower magnitude values after binary representation of individual digits will be (0.0001),(0),(0),(0101) and (0.0101) respectively.

Step 2: Repeat step 1 until all magnitude values of selected harmonic values are converted into set of small magnitude values.

Therefore, if N is the size of harmonic component, the size of lower magnitude values would be 5 X N.

2.3. Embedding Secret Code (GASASH-ESC)

The method of embedding secret code with original song is described with the help of wavelet transform is fabricated with embedding secret code with coefficient components of the amplitude signal without affecting its audible quality. The procedure of generating secret signal is depicted in the following algorithm.

Algorithm:

Input: Original song signal and frequency band (Val) [generated by section 2.2].

Output: Modified song signal with embedding secret code.

Method: Separating amplitude and phase of original song and applying single-level discrete 2-D wavelet transform [symmetrization mode] are described in the following steps.

Step 1: Separating phase and amplitude signal from original song.

Step 2: Apply single-level discrete 2-D wavelet transform [symmetrization mode] over amplitude and phase signals to generate the approximation coefficients matrix approximation coefficients (CA) and details coefficients matrices horizontal detail coefficients (CH), vertical detail coefficients (CV), diagonal detail coefficients (CD) for each of them. The mathematical expressions of the wavelet transform are as follows.

A signal, f(t), can be represent using wavelet transform as given in equation 1.

$$f(t) = \sum_m \sum_n C_{m,n} \psi_{m,n}(t) \quad (1)$$

Where both m and n are integer and $\psi_{m,n}(t)$ represents successive scaled and dilated versions of a single wavelet function $\psi(t)$ [mother wavelet] [5]. The wavelet basis functions can be formulated with the help of mother wavelet as given in equation (2).

$$\psi_{m,n}(t) = 2^{-m/2} \psi(2^{-m}t - n) \quad (2)$$

The above factor $2^{-m/2}$ usually maintains a constant relationship, i.e., the wavelet basis is independent of scale m and represents the successive partitioning of signal's spectrum throughout wavelet scaling and dilations. In case of iterative wavelet transform, real-valued coefficients, i.e.,

$\alpha_{m,n}$ are used. The approximation and wavelet coefficients $c_{m,n}$ and $\alpha_{m,n}$ can be represented as follows:

$$\alpha_{m,n}(t) = \sum_k h_{2^{n-k}} \alpha_{m-1,k} \quad (3)$$

$$c_{m,n}(t) = \sum_k g_{2^{n-k}} c_{m-1,k} \quad (4)$$

Where h and g are low pass FIR and high pass FIR filters respectively [7]. For the 2-D DWT separate horizontal and vertical audio signal filtering and downsampling are needed. The 2-D DWT produces four sub-bands at each level of analysis. The first is the scaling of the input audio signal (CA) and the remaining three are the detail coefficients at the horizontal (CH), vertical (CV) and diagonal (CD) directions respectively.

Step 3: It is experimentally observed the values of diagonal detail coefficients (CD) for both amplitude and phase signals of original song are less difference. If we make the both values of CDs equal then, no affect on audible quality has been proved [8]. If the difference values of two CDs become higher for a particular audio, then more higher level discrete 2-D wavelet transform in symmetrization mode may be applied as required.

Step 4: Embedding lower magnitude (Val) values with diagonal detail coefficients (CD) of amplitude signal as given below.

- i. Add i^{th} sampled value ,Val, with k^{th} position of CD as follows

$$CD[k,C] = CD[k,C] + Val [i] , \quad \text{where } C \text{ indicates the channel number,}$$

In case of mono type song C is not needed
- ii. Add next value of Val after n positions from k^{th} position of the alternative channel (if previous inserted value at C=1, then C=2, or vice versa) of CD, where n represents gaps between two consecutive embedded positions. The value of n is after adding 0 is 1, after 0.0001 is 2, after 0.0010 is 3, ... and after 0.1001 is 10. 2 additional gaps will added when all parts of magnitude values of a particular magnitude value is appended with CD for separating from next intensity value as well as another 2 gaps will be added after inserting all individual digits of each magnitude value of secret code.
- iii. Repeat above step i until all values of Val are not added with CD.

Step 5: Reconstruct the both amplitude and phase signal using the inverse DWT out of calculated wavelet and scaling coefficients and generate song signal with reconstructed amplitude and song signals.

Though the inserted magnitude values of Val (of selected frequency band) are very small, therefore, appending extra values with original song (cover song) will not affect its overall audible quality but able to separate the original song from the similar songs.

2.4. Extraction

The decoding is performed using similar mathematical calculations. The algorithm of the same is given below.

Algorithm:

Input: Modified song signal with embedded secret code.

Output: Original song signal.

Method: The details of extraction of original song signal are given below.

Step 1: Separate amplitude and phase from song signal and find the diagonal detail coefficients (CD) as step 1 and 2 of embedding secret code [section 2.2].

Step 2: if 1 gap found between unequal values of two CDs then added digit is 0 , if n gaps found then digit is (n+1), where $0 \leq n \leq 9$, if additional 2 gaps found, add the all previous digits to get the individual digit of intensity value. Again additional gaps 2 found then put previous individual digit side by side to get the final magnitude value as represented by following example. Let, generated individual digits for a particular V are 1, 0, 0, 5 and 1[in sequence]. Then corresponding magnitude value would be 1.0051.

Step 3: Deduct the extra value from particular channel of CD [extra value should be added or subtracted based on negative or positive magnitude value of the channel] of amplitude signal comparing with CD of phase signal.

Step 4: Repeat step 2-3 until all extra values are deducted from CD's component of amplitude signal.

Step 5: Reconstruct the both amplitude and phase signal using the inverse DWT out of calculated wavelet and scaling coefficients and generate song signal with reconstructed amplitude and song signals.

3. EXPERIMENTAL RESULTS

Encoding and decoding technique have been applied over 1 minute recorded songs, the song is represented by complete procedure along with results in each intermediate step has been outlined in subsections 3.1.1 to 3.1.4. The results are discussed in two sections out of which 3.1 deals with result associated with GASASH and that of 3.2 gives a comparison with existing techniques.

3.1 Results

For experimental observation, a strip of 1 minute song ('The Catalyst', sang by Linkin Park) has been taken. Figure 1 shows amplitude-time graph of the original signal. Figure 2 is representing the concealed secret code. GASASH is applied on this signal and the output generated in the process is shown in figure 3. Figure 4 shows the difference of amplitude values before and after modification of original song. From figure 4 it is seen that the deviation of the modified song signal is very less, i.e., its audible quality will not be affected at all.

3.1.1. Original Recorded Song Signal (1 minute)

The graphical representation of the original song, considering sampled values (441000) of $x(n,2)$ [stereo type song] is given in the figure 1(a). Figure 1(b) and 1(c) are shown amplitude and phase signals respectively.

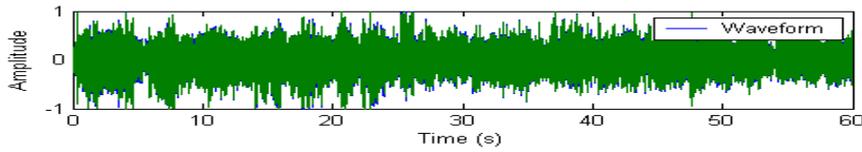


Figure 1(a). Original song ('The Catalyst', sang by Linkin Park)

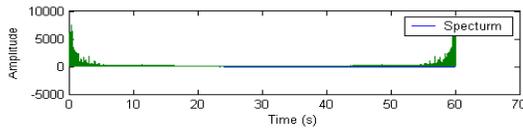


Figure 1(b). Amplitude signal

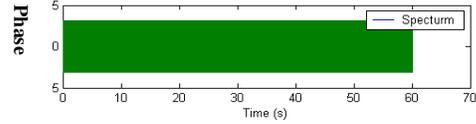


Figure 1(c). Phase signal

3.1.2. Selected Hidden Frequencies

The concealed frequency band of song signal is shown in figure 2.

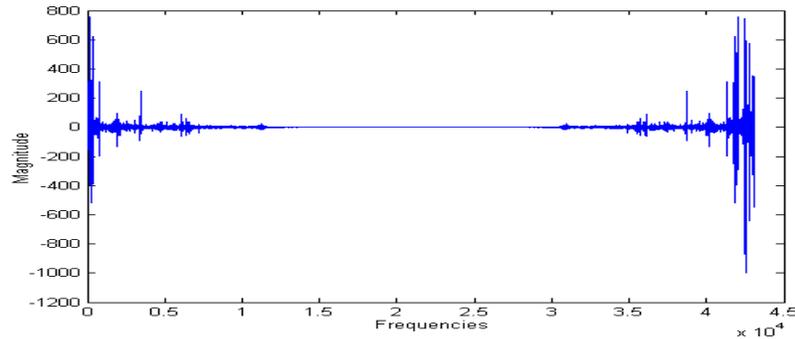


Figure 2 Selected frequencies

3.1.3. Modified song after embedding secret code (1 minute)

The graphical representation of the modified authenticated song signal is shown in the figure 3.

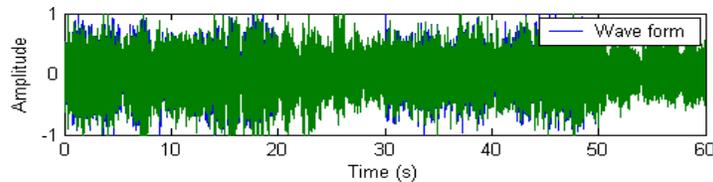


Figure 3 Modified song with embedded secret code

3.1.4. The difference of magnitude values between original and modified songs

The graphical representation of the difference of magnitude values of original and modified songs is shown in the figure 4.

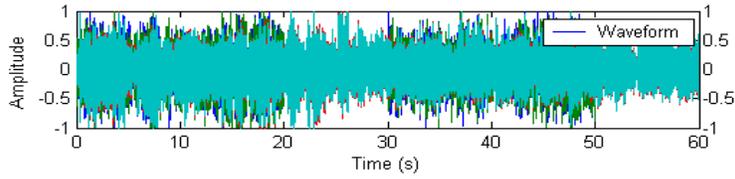


Figure 4: The difference of sampled values of two signals as shown in figure 1 and 2.

3.2 Comparison with existing systems

Various algorithms [5] are available for embedding information with audio signals. They usually do not care about the quality of audio but we are enforcing our authentication technique without changing the quality of song. A comparison study of properties of our proposed method with Data Hiding via Phase Manipulation of Audio Signals (DHPMA)[3] and Secret Image Embedded Authentication of Song Signal through Wavelet Transform (IAWT)[9] before and after embedding secret message/modifying parts of signal (16-bit stereo audio signals sampled at 44.1 kHz) is given in table 1, table2 and table3. Average absolute difference (AD) is used as the dissimilarity measurement between original song and modified song to justify the modified song. Whereas a lower value of AD signifies lesser error in the modified song. Normalized average absolute difference (NAD) is quantization error is to measure normalized distance to a range between 0 and 1. Mean square error (MSE) is the cumulative squared error between the embedded song and the original song. A lower value of MSE signifies lesser error in the embedded song. The SNR is used to measure how much a signal has been tainted by noise. It represents embedding errors between original song and modified song and calculated as the ratio of signal power (original song) to the noise power corrupting the signal. A ratio higher than 1:1 indicates more signal than noise. The PSNR is often used to assess the quality measurement between the original and a modified song. The higher the PSNR represents the better the quality of the modified song. Thus from our experimental results of benchmarking parameters (NAD, MSE, NMSE, SNR and PSNR) in proposed method obtain better performances without affecting the audio quality of song.

Table 1. Metric for different distortions

Sl No	Statistical parameters for differential distortion	Value using GASASH	Value using IAWT [9]	Value using DHPMA [3]
1	MD	0.1512	0.1431	0.1854
2	AD	0.0423	0.032	0.0294
3	NAD	0.1563	0.1421	0.1910
4	MSE	8.263e-004	8.341e-004	0.0017
5	NMSE	0.901	0.0835	1.5271e-005

Table 2 gives the experimental results in terms of SNR (Signal to Noise Ratio) and PSNR (Peak signal to Noise Ratio). Table 3 represents comparative values of Normalized Cross-Correlation (NC) and Correlation Quality (QC) of proposed algorithm with DHPMA.

Table 2. SNR and PSNR

Sl No	Statistical parameters for differential distortion	Value using GASASH	Value using IAWT [9]	Value using DHPMA [3]
1	Signal to Noise Ratio (SNR)	17.3211	18.4331	20.0130
2	Peak Signal to Noise Ratio (PSNR)	33.7522	31.5487	27.5231

Table 3. Representing NC and QC

Sl No	Statistical parameters for correlation distortion	Value using GASASH	Value using IAWT [9]	Value using DHPMA [3]
1	Normalised Cross-Correlation (NC)	1	1	1
2	Correlation Quality (QC)	-0.0734	-0.0719	-0.0721

The Table 4 shows PSNR, SNR, BER (Bit Error Rate) and MOS (Mean opinion score) values for the proposed algorithm. Here all the BER values are 0. The figure 4 summarizes the results of this experimental test. It shows this algorithm’s performance is stable for different types of audio signals.

Table 4. Showing SNR, PSNR BER, MOS

Audio (1s)	SNR	PSNR	BER	MOS
Song1	17.3211	33.7522	0	5
Song2	15.4421	32.753	0	5
Song3	17.216	30.8011	0	5
Song4	20.331	36.2651	0	5
Song5	16.3441	30.9131	0	5

This quality rating (Mean opinion score) is computed by using equation (5).

$$Quality = \frac{5}{1 + N * SNR} \tag{5}$$

Where N is a normalization constant and SNR is the measured signal to noise ratio. The ITU-R Rec. 500 quality rating is perfectly suited for this task, as it gives a quality rating on a scale of 1 to 5 [6]. Table 5 shows the rating scale, along with the quality level being represented.

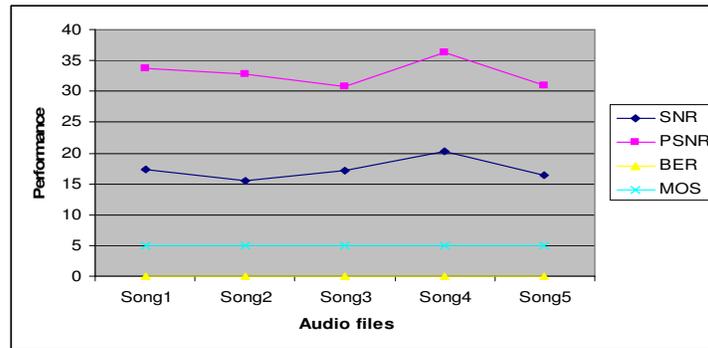


Figure 4. Performance for different audio signals.

Table 5. Quality rating scale

Rating	Impairment	Quality
5	Imperceptible	Excellent
4	Perceptible, not annoying	Good
3	Slightly annoying	Fair
2	Annoying	Poor
1	Very annoying	Bad

3. CONCLUSIONS AND FUTURE WORKS

In this paper, an algorithm for generating the hidden authenticating code with the help of wavelet transform for selected coefficient of song signal and embedding secret image in the specified region of that coefficient has been proposed which will not affect the song quality but it will ensure to detect the distortion of song signal characteristics. Finding secret frequency band and comparing with similar existing inner frequency band of song signal, easily determine the original one from collection of similar versions. The modified song with authenticated secret code will not affect the song quality but ensure to detect the distortion of song signal characteristics.

This technique is developed based on the observation of characteristics of different songs but the mathematical model for representing the variation of those characteristics after modification may be formulated in future. It also can be extended to embed an audio into song signal instead of frequency components or numeric values. The perfect estimation of percentage of threshold numbers of sample data of song that can be allow to change for a normal conditions will be done in future with all possibilities of errors.

ACKNOWLEDGEMENTS

The authors express gratitude to the Dept. of CSE & IT, College of Engg. & Management, Kolaghat and Dept of CSE and the PURSE scheme of DST, Govt. of India of University of Kalyani, under which the research has been carried out.

REFERENCES

- [1] Mondal, Uttam Kr., Mandal, J.K.: A Practical Approach of Embedding Secret Key to Authenticate Tagore Songs(ESKATS), Wireless Information Networks & Business Information System Proceedings (WINBIS'10), ISSN 2091-0266,organized by Rural Nepal Technical Academy (Pvt.) Ltd , Nepal, Volume 6, Number 1, pp 67-74(2010).
- [2] Mondal, Uttam Kr., Mandal, J.K. : A Novel Technique to Protect Piracy of Quality Songs through Amplitude Manipulation (PPAM), International Symposium on Electronic System Design (ISED 2010), ISBN 978-0-7695-4294-2 ,pp 246-250(2010).
- [3] Xiaoxiao, Dong, Mark, F., Bocko, Zeljko Ignjatovic: Data Hiding Via Phase Manipulation of Audio Signals , IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2004) , ISBN 0-7803-8484-9 ,Vol 5, pp 377-380(2004).
- [4] Erten, G., Salam, F.:Voice Output Extraction by Signal Separation ” , ISCAS '98 ,ISBN 07803-4455-3,Vol 3, pp 5 - 8(1998).
- [5] Pohl, C., Genderen, J.L. Van : “Multisensor image fusion in remote sensing concepts, methods and applications” International journal of remote sensing, Vol. 19, No.5, pp. 823-854.
- [6] Arnold, M.: Audio watermarking: Features, applications and algorithms, IEEE International Conference on Multimedia and Expo ,New York, NY, Vol 2, pp. 1013-1016, (2000).
- [7] Ioannidou, Stella, Karathanassi, V., Sarris, A.: “The optimum wavelet-based fusion method for urban area mapping”, WSEAS International Conference on Environment, Ecosystems and Development.
- [8] Mondal, Uttam Kr., Mandal, J.K.: “ Generation and Fabrication of Authenticated Song Signal through Preserving Balance Ratio of Song's Components (BRSAS)”, International Conference on Communication and Industrial Application (ICCIA 2011), Kolkata, India, 26th- 28th Dec,2011.
- [9] Mondal, Uttam Kr., Mandal, J.K.: “Secret Image Embedded Authentication of Song Signal through Wavelet Transform (IAWT)”, The Fifth International Conference on Network Security & Applications (CNSA 2012), Chennai, India, Volume 176, 199-208, 2012.

Authors

Uttam Kr. Mondal, has received his Bachelor of Engineering (B.E) degree in Information Technology in 2004 and Master of Technology (M.Tech) in Information Technology in 2006 from University of Calcutta, India. He has now working as an Asst. Professor in department of Computer Science & Engineering and Information Technology in College of Engg. & Management, Kalaghat, West Bengal, India. His research areas include cryptography & Network Security, Audio signal authentication. He has 15 publications in National and International conference proceedings and journal.



Jyotsna Kumar Mandal, M. Tech.(Computer Science, University of Calcutta),Ph.D.(Engg., Jadavpur University) in the field of Data Compression and Error Correction Techniques, Professor in Computer Science and Engineering, University of Kalyani, India. Life Member of Computer Society of India since 1992 and life member of cryptology Research Society of India. Dean Faculty of Engineering, Technology & Management, working in the field of Network Security, Steganography, Remote Sensing & GIS Application, Image Processing. 25 years of teaching and research experiences. Eight Scholars awarded Ph.D. one submitted and eight are pursuing. Total number of publications is more than two hundred thirty in addition of publication of five books from LAP Lambert, Germany.

