I.J. Image, Graphics and Signal Processing, **2013, 3, 42-47**
Published Online March 2013 in MECS (http://www.mecs-press.org/)
DOI: 10.5815/ijigsp.2013.03.06

A Unique Wavelet Steganography Based Voice Biometric Protection Scheme

Dr Sanjaypande M. B
Prof. and Head of Dept.,VVIET, Mysore, India
E-mail: rkroop99@gmail.com

Raikoti Sharanabasappa
Research Scholar, India
E-mail: sr.raikoti@gmail.com

***Abstract*—** Voice biometric is an easy and cost effective biometric technique which requires minimalistic hardware and software complexity. General voice biometric needs a voice phrase by user which is processed with Mel Filter and Vector Quantized features are extracted. Vector quantization reduces the codebook size but decreases the accuracy of recognition. Therefore we propose a voice biometric system where voice file's non quantized code books are matched with spoken phrase. In order to ensure security to such direct voice sample we embed the voice file in a randomly selected image using DWT technique. Imposters are exposed to only images and are unaware of the voice files. We show that the technique produces better efficiency in comparison to VQ based technique.

***Index Terms*—** Image Steganography, Audio Steganography, Discrete Wavelet Transform, Spectrum Analysis, Voice Biometric, Voice Recognition

I. INTRODUCTION

Much research has been done in the area of speaker verification using cepstral analysis, end point detection algorithms, pattern recognition, neural networks, stochastic models and many-distance measuring algorithms. However as per media reports, the satisfactory performance of the existing systems is still a matter of great concern because of the considerable number of false acceptances and false rejections.

The Speaker Verification can be used in many areas. Some of them are Access control to computers, cellular phones, databases, Access control to professional or wide public sites, Protection of confidential data, Remote access to computer networks, Electronic commerce, Forensic, Automatic door opening systems when person arrives, Telephone-banking transactions and Voice commands on cellular phones. Many organizations like banks, institutions, industries etc are currently using this technology for providing greater security to their vast databases.

Given the voice input, the objective of the word-dependent Voice recognition system is to verify whether the speaker is who he/she claims to be and whether he has spoken his key phrase or not . The system processes voice signal of the user, which is given as input and finds the Mel Frequency Cepstrum Coefficients. Then it generates a template, which is called a codebook, an array of acoustic vectors. In the enrollment phase the user codebook is saved in the system. In the verification phase the user codebook is compared against the claimed user's actual codebook, which is stored in the system during the enrollment phase. If the difference is below the threshold value the user is authenticated else the user is not authenticated.

During the first phase, speaker enrollment as shown in the Fig. 1 (a), features are extracted from the input speech signal given by the speaker by a process called Feature extraction, and are modeled as a template. The modeling is a process of enrolling speaker to the verification system by constructing a model of his/her voice, based on the features extracted from his/her speech sample. The collection of all such enrolled models is called speaker database.

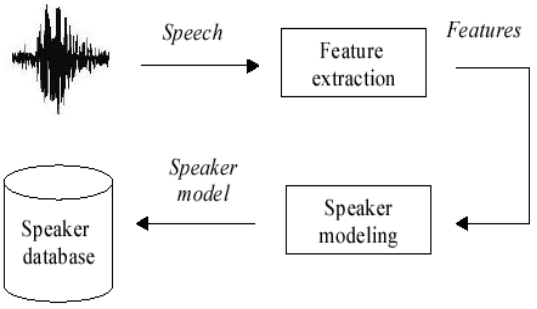


Fig. 1(a): Schematic flow of Enrollment Phase

In the second phase, verification phase as shown in the Fig.1. (b), features are extracted from the speech signal of a speaker and these current features are compared with the claimed features stored in the database by a process called Feature matching. Based on this comparison the final decision is made about the speaker identity.

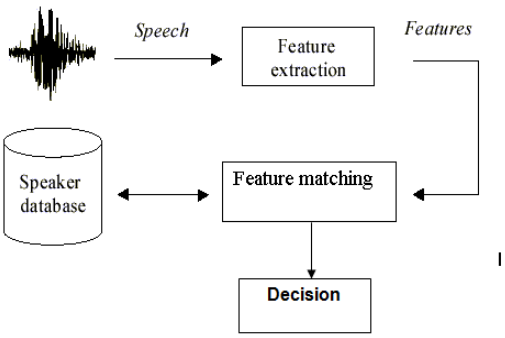


Fig. 1 (b): Schematic flow of Verification Phase

Further, the feature extraction process is carried out as extraction of significant frequency components from voice file. Firstly, the voice sample is processed and silence period is removed, followed by filtering and windowing the voice sample. This windowed result is quantized to generate templates.

Once templates are generated, test template is checked with trained templates for matching with closest distance matching. The process is depicted in figure 2.

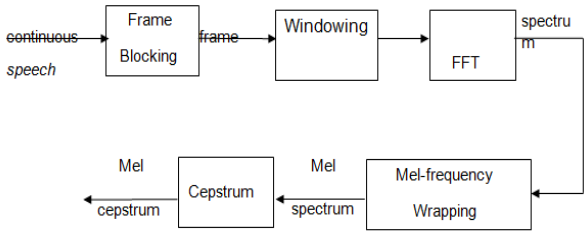


Fig. 2: Block diagram of the MFCC processor

However reverse engineering the process is comparatively easy by the imposter. All, he has to do to self authenticate is , create his own vector quantized voice sample of any spoken phrase and change any of the database phases to this value. If database changes are not allowed than he can simply try generating a phrase to get VQ values as close to a selected trained VQ file.

Therefore security must be adopted to ensure that imposters does not have a direct access to the feature vectors. Several template protection schemes have been proposed in the past for template security. But as voice templates in many ways are different from other templates (as voice is audible while supplying to testing process) , a more sophisticated and technique is needed.

In this work we introduce a novel technique of comparing codebook which is magnitude features from MEL cepstrum rather than passing it through vector quantization stage.

In order to provide more security we use image steganography. Steganography change the basic form of a data which hides the data from any imposter. Therefore it provide high level of reliability to the method.

We define the important terminology as bellow.

II. Related Work

Souvik Bhattacharyya and Gautam Sanyal[1] proposed approach works by converting the gray level image in transform domain using discrete integer wavelet technique through lifting scheme. The aim of this paper is to propose a high-capacity image steganography technique that uses pixel mapping method in integer wavelet domain with acceptable levels of imperceptibility and distortion in the cover image and high level of overall security.

Ammar Abdul-Amer Rashed[2] proposed a companied technique for hiding secret messages (text) based on wavelet transform applying in cover image (a gray level image 8bit) and Huffman encoding. The experimental results show that the algorithm has a high capacity and a good invisibility, Moreover PSNR of stego image shows the better results the PSNR above 40 dB, the proposal system was activated according to attacker noise is addition and JPEG compression application are used without detection the secret message.

S. K. Muttoo and Sushil Kumar[3] proposed a stiganographic algorithm based on wavelet transforms. The algorithm first uses the Best T-codes to encode the message before embedding into a cover image..

Saddaf Rubab and Dr. M. Younus presented[4] a new devised algorithm to hide text in any colored image of any size using Huffman encryption and 2D Wavelet Transform. The subject algorithm also proved secure as Huffman table is required to decode the information.

Manjunatha Reddy and Raja proposed[5] High Capacity and Security Steganography using discrete wavelet transform (HCSSD). The wavelet coefficients of both the cover and payload are fused into single image using embedding strength parameters alpha and beta. The cover and payload are preprocessed to reduce the pixel range to ensure the payload is recovered accurately at the destination. It is observed that the capacity and security is increased with acceptable PSNR in the proposed algorithm compared to the existing algorithms.

Lalitha.G et al.[6], proposed a technique for the simultaneous transmission of multiple data securely. They took an advantage of less space required for storing an image than that of a wav file. The proposed technique brings down the required channel capacity to transfer secret data in real time systems besides improving robustness.

Elham Ghasemi et al.[7], proposed the application of Wavelet Transform and Genetic Algorithm in a novel steganography scheme. We employ a genetic algorithm based mapping function to embed data in Discrete Wavelet Transform coefficients in 4x4 blocks on the cover image.

Kirith saroha and Pradeep kumar singh[8] proposed a new steganographic method for embedding an image in an Audio file. Emphasis will be on the proposed scheme of image hiding in audio and its comparison with simple Least Significant Bit insertion method of data hiding in audio. It is an attempt to find a method that uses an audio file as a cover media to hide an image without making noticeable changes to the file structure and contents of the audio file. The proposed scheme is based on Least Significant Bit insertion method as it has been already proved that modification of LSB creates a minimal change in the audio file format.

Akram M. Zeki et al[9]., provided analysis on steganographic techniques and undertake an experiment using five Steganographic software in order to explore their capabilities. Benchmarking tool for identifying different performance aspects of the Steganographic techniques and Steganographic software like visual quality, performance indices, memory requirement and the evaluation of the maximum capacity for each software under this study.

Jayaram P et al[10]., made a survey on audio steganography. They proposed that Least Significant Bit (LSB) coding method is the simplest way to embed secret information in a digital audio file. This proposed method provides greater security and it is an efficient method for hiding the secret information from hackers and sent to the destination in a safe and undetectable manner. This proposed system also ensures that the size of the file is not changed even after encoding and it is also suitable for any type of audio file format.

Md. Shafakhatullah Khan et al.[11], proposed a new approach which is sophisticated for concealing the data. They describes how the data is secured form the intruders even though they trace the audio file which contains the confidential data.The basic idea is to provide an optimized method for concealing the private data from intruders and sent to the destination in a safer and secure manner. It is an enhancement of spread spectrum audio data hiding methods.

Wei Qin Cheng et al.[12], proposed a robust audio steganography. They implemented a simple Dynamic Linked Library [DLL] by using managed C++ and Microsoft .NET framework. It is implemented by Direct Sequence Spread Spectrum [DSSS] method on data block base.

[13] and [14] provides an overview of voice properties and information about biometrics and speech technology.The information needed for the methodology, which involves Mel Frequency Cepstrum Coefficients technique for Feature extraction process and Vector Quantization technique for Feature matching process, is taken from [13] and [14]. Feature extraction is the process that extracts a small amount of data from the voice signal that can later be used to represent the speaker. Feature matching involves the actual procedure to verify the speaker by comparing extracted features from his/her voice input with the claimed one which is stored in the database.

III. Problem Definition

Several works have been presented towards both steganography, cryptography as well as combining the techniques for better data security.

Even though the research of Steganography initially was carried out as a problem of hiding data behind cover images, as the technology has evolved more complex forms of steganography has eveolved.

Some of the widely researched and evolved variants of Steganography are:

1) Hiding Image Behind Video
2) Hiding Data Behind Video
3) Hiding data behind Audio
4) Hiding Audio Behind Video

There have been little effort to use steganography to replace conventional template security techniques. Biometric is a growing field and demands every possible security extension for data and biometric template security.

However many reversible and irreversible methods for template security have been proposed. But as discussed earlier, unlike other biometric forms like face and fingerprint biometric , voices are easy to imitate. Therefore mere cryptographic use for securing voice templates are not sufficient.

Therefore we change the mode of processing by first storing the voice itself behind an image to change the nature of file.

Templates from training instances are generated at the run time by first extracting the voices from image followed by template extraction. As templates are extracted and matched at the run time, possibility of tempering is reduced significantly.

Hence we propose voice biometric template security by hiding the recorded user voice files behind random images.

IV. Proposed System

Encoding and Decoding techniques are clearly presented through figure 3 and 4 respectively.

First a cover image is selected randomly from database for hiding the trained file of a user. Image is decomposed through Haar wavelet to produce multi scale image. High pixel density areas are identified. Audio file is normalized to fit the data range of this file. the file is then hidden behind the image by storing normalized audio bits behind wavelet image bits. Inverse transform is applied and actual image is recovered.

In the decoding process the cover image is transformed through wavelet. First dense regions are

identified. Than normalized voice data is extracted and remapped to original scale. The data is then processed as explained in figure 3 for generating templates.

Encoding Algorithm

1. Read a Gray scale Image
2. Take Wavelt Transform of the Image
3. Find the maximum of the data and Normalize the data by dividing it with MAx
4. Hide the normalized data in cV and cD
5. Reconstruct the image from wavelet
6. Reshape the image to single dimensional array.
7. Take the spectrum of the signal. (>3.3KHz)
8. Store the length
9. Store reshaped audio
10. Coverst 1d to 2D

Decoding Algorithm

1. Perform Inverse DWT
2. Extract the audio length
3. Extract Normalized audio
4. Denormalization
5. Extract voice data
6. Process for Recognition

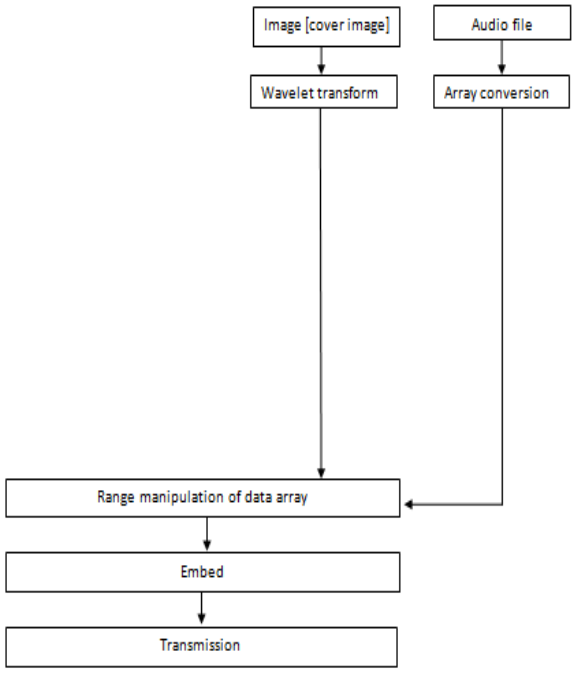


Figure 3:Encoding stage Block Diagram

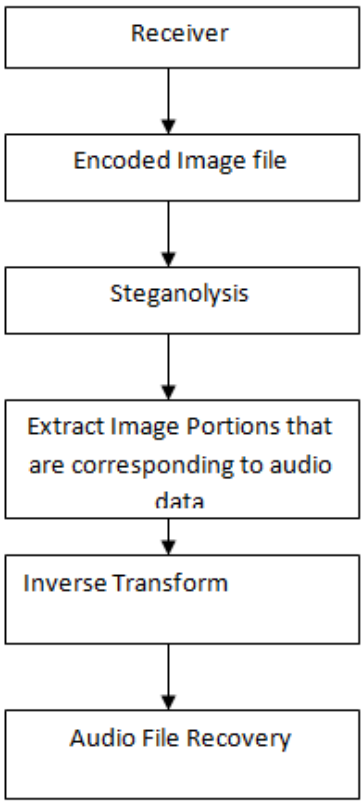


Figure 4: Decoding stage block diagram

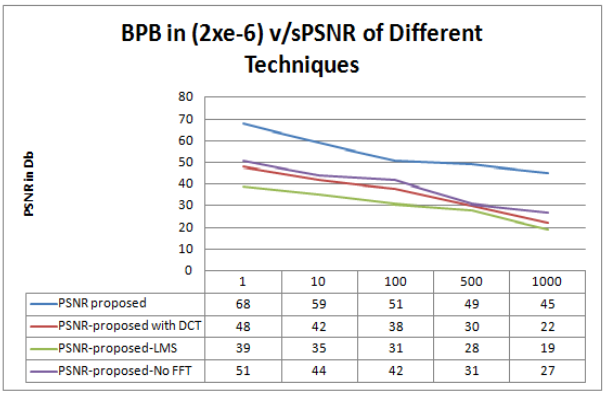
V. RESULTS AND ANALYSIS

As there are not many variants of Hybrid Steganography, We have changed the core model of the work and have compared the performances. The proposed work of DWT based image steganography followed by Spectrum based Audio steganography is compared will following other approaches.

a) DCT based Image steganography, followed by spectrum based audio steganography
b) LMS image steganography followed by Specturn based digital Steganography
c) DWT based image Steganography with LMS based Audio Steganography.

BPP is changed by varying the payload bits. All the experiments are conducted for Uncompressed Monochrome Image of 256x256 size and wav audio file of 2Mb. BPB is measured as Number bits of payload hidden par bit of audio file.

Results are shown in Figure 5.1. Results show that the proposed technique is a clear winner in terms of performances against all other forms of steganography compared here.



	1	10	100	500	1000
PSNR proposed	68	59	51	49	45
PSNR-proposed with DCT	48	42	38	30	22
PSNR-proposed-LMS	39	35	31	28	19
PSNR-proposed-No FFT	51	44	42	31	27

Figure 5: Performance analysis of proposed system with other conventional techniques.

VI. Conclusions and Prospects

Security plays a vital role in all aspects of Biometric. As the technology grows, its bounds in the positive direction, so do the reverse engineering of it. Data security has been of chief concern these days and plays a major role in terms of its complexity. One such means of providing data security is using steganography.

The proposed work is one of the latest advancements accomplished in the field of biometric security. In this Work a user's voice data known as the payload is embedded into an image which is called the cover image. For a general view it seems as a simple Image file.

A question now arises as to how this will be secure as one can easily identify the distortions in the image.

The BPP analysis proves that the Bits par pixel of the technique is very high. Therefore image distortions are minimum.

Audio quality to be not that good and can try to decode. Performance analysis of the recognition accuracy shows that the proposed system produces a voice recognition accurac y of 88% with only .3% false acceptance rate as against VQ based technique which produces a overall accuracy of 81% with 6% FA.

References

[1] K B Shiva Kumar et. al. "Bit length replacement steganography based on DCT coefficients" / International Journal of Engineering Science and Technology. Vol. 2(8), 2010, 3561-3570

[2] K B Shiva Kumar et. al "Hybrid Domain in LSB Steganography" International Journal of Computer Applications (0975 – 8887). Volume 19– No.7, April 2011

[3] K B Shiva Kumar et. al "Steganography Based on Payload Transformation" IJCSI International Journal of Computer Science Issues, Vol. 8, Issue 2, March 2011. ISSN (Online): 1694-0814. www.IJCSI.org

[4] Saddaf Rubab, Dr. M. Younus. "Improved Image Steganography Technique for Colored Images using Huffman Encoding with Symlet Wavelets" IJCSI International Journal of Computer Science Issues, Vol. 9, Issue 2, No 1, March 2012 ISSN (Online): 1694-0814".

[5] Akram M. Zeki, Adamu A. Ibrahim And Azizah A. Manaf, "Steganographic Software: Analysis and Implementation" International Journal Of Computers And Communications Issue 1, Volume 6, 2012

[6] S. K. Muttoo and Sushil Kumar, "Robust Source Coding Steganographic Technique Using Wavelet Transforms" BVICAM's International Journal of Information Technology.

[7] Kriti Saroha and Pradeep Kumar Singh "A Variant of LSB Steganography for Hiding Images in Audio". International Journal of Computer Applications (0975 – 8887) Volume 11– No.6, December 2

[8] Lalitha.G et al. / International Journal on Computer Science and Engineering (IJCSE), "Secure Transmission of Compound Information Using Image Steganography"

[9] Md. Shafakhatullah Khan et al. "An Optimized Method for Concealing Data using Audio Steganography" International Journal of Computer Applications (0975 – 8887) Volume 33– No.4, November 2011

[10] Pradeep Kumar Singh et. al "Enhancement of LSB based Steganography for Hiding Image in Audio". / (IJCSE) International Journal on Computer Science and Engineering Vol. 02, No. 05, 2010, 1652-1658

[11] Jayaram et al. "Information Hiding Using Audio Steganography –A Survey"The International Journal of Multimedia & Its Applications (IJMA) Vol.3, No.3, August 2011

[12] Elham Ghasemi, Jamshid Shanbehzadeh, Nima Fassihi. " High Capacity Image Steganography using Wavelet Transform and Genetic Algorithm". Proceedings of the international multiconference of engineers and computer scientists 2011 volI

[13] Minh N.Do, *An Automatic speaker Recognition System*, Audio Visual Communications Laboratory, Swiss Federal Institute of Technology, Lausanne, Switzarland.

[14] J.W. Cooley and J.W. Tukey, *An algorithm for the machine calculation of complex Fourier Series*, Mathematics Computation, Vol.19, 1965, pp 297-301.

S. R Raikoti is M.Tech and M. Phill in Computer Science. He is currently pursuing research in Biometric Imaging. He has guided over 50 M. Phill students in their research work. His area of interest are Image Processing, Pattern Recognition, Biometric templates and Template security.

Dr. Sanjay Pande M. B. is professor and Head of Department of Computer Science Department in VVIET, Mysore. He received his P.H.D from Mysore University in 2005. His area of interest is Imaging studies for cognition and recognition in Pattern Recognition.

