

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 7, Issue 2, February 2018

A Virtual Reality Based 3-Tier Secure Speech Communication for Defense Applications

G. Ridhu¹, V.Saveetha, M. Tech., MBA²,

M.E. (Communication and Networking), Department of Electronics and Communication Engineering, Pavendar
Bharathidasan College of Engineering and Technology, Mathur, Thiruchirappalli, India

Assistant Professor, Department of Electronics and Communication Engineering, Pavendar Bharathidasan College of
Engineering and Technology, Mathur, Thiruchirappalli, India

ABSTRACT: Advanced Digital Speech Recordings seem imperative in everyday life, making the requirement for important scientific technologies. A standout amongst the most imperative issues is the recognition of twofold compression since it frequently takes after data/content misuses in the exchange of information. To keep up the privacy of such data it is important to encode and encrypt it. The most imperative and productive Speech Communication is fundamental in numerous fields of military exercises. In this system, a new procedure is introduced for distinguishing twofold compacted "AMR" sound utilized as a part of military exchanges. For a voice summon with numerous edges, the highlights of the considerable number of edges are totaled and ordered. Another procedure for Speech Compression and encryption alongside covering up of the information is presented. Speech Compression changes over human speech signals into a minimal shape which ends up noticeably solid for communication and capacity by decreasing the extent of information without losing nature of the first speech. Speech encryption changes over the ordinary type of speeches into unrecognizable shape which builds the security of correspondence through a shaky channel. In this manner, presentation of Discrete Wavelet Transformation [DWT] for twofold compression gives more exactness and secure correspondence for military events. For all the advanced speech processing based on encryption norms are more secured and reliable over communication field especially in secured environments such as military, navy and so on. The entire system operation is handled via new technological invention called Virtual Reality [VR], which enables the system to generate the key for cryptology, according to that the data is converted to cipher format and the resulting accuracy of virtual reality and its processing is maximum compare to existing strategies.

KEYWORDS: Speech Analysis, Information Security, Speech Compression, Encryption, DWT, Discrete Wavelet Transformation, Virtual Reality, VR.

I. INTRODUCTION

Productive Speech Communication is perceived as a basic and instrumental capacity in numerous fields of military events, for example, charge/command and control, airforce, navy activities, military interchanges, interpretation, insight, and preparing. Speech Coding calculations are utilized as a part of low-bit-rate military voice communication frameworks. Speech info and yield frameworks can be utilized as a part of summon and control situations to considerably diminish the workload of administrators. By and large administrators have occupied eyes and hands, and should utilize different intends to trigger control works and get input data. Speech handling strategies are accessible to distinguish talkers, dialects, and catchphrases and can be coordinated into military knowledge frameworks. PC helped preparing frameworks consolidating programmed Speech acknowledgment and combination advances can be used to prepare work force with least or no instructor-bolster. Interchanges, command and control, insight and preparing frameworks are making increasingly utilization of Speech and dialect innovation segments: that is, Speech coders, Voice Controlled C2 frameworks, speaker and dialect acknowledgment, interpretation frameworks and mechanized preparing suites. Present day Speech innovation depends on advanced flag preparing, probabilistic

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 7, Issue 2, February 2018

hypothesis and hunt calculations. These procedures make it conceivable to perform noteworthy information diminishment for coding and transmission of Speech signals, Speech union and programmed acknowledgment of Speech, speaker or dialect. By using the concept of Virtual Reality [VR], we can make more security compare to other process for generating and manipulating keys.

II. SPEECH COMPRESSION

The AMR audio compression codec is enhanced for speech coding. A run of the misuses of AMR sound includes decompressing the AMR sound into a waveform, altering the flag by sound altering programming and re-compressing it into the AMR arrange. In this manner, the altered sound is compacted twice, and is regularly alluded to as twofold compressed AMR sound. The genuineness of twofold compressed sound is sketchy in light of the fact that it might have been already changed. In this framework, the issue of recognizing twofold compressed AMR sound is tended to. As of late, specialists have tended to numerous sound legal issues, for example, altering discovery, recorder distinguishing proof, speaker ID and check, and audio compression history investigation. Pressure history investigation is useful in computerized sound legal sciences and has drawn the consideration of scientists in the previous decade. The primary thought behind recognizing a audio compression history is to dissect the presented quantization antiques. Yang et al. proposed a strategy to uncover counterfeit quality MP3 sound clasps re-compressed with higher piece rates (up-transcoding).

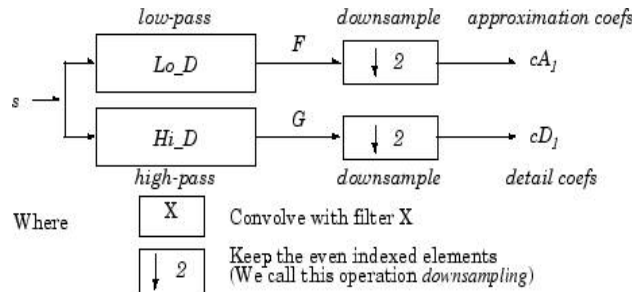


Fig.1 DWT Processing

They examined the MDCT (altered discrete cosine change) coefficients in MP3 sound and discovered that the quantity of MDCT coefficients with little esteems in counterfeit quality MP3 is lower than typical MP3. In view of twofold quantization ancient rarities, a few scientists utilized factual highlights of MDCT coefficients to recognize MP3 re-compacted cuts for both up-transcoding and down-transcoding Jenner et al. proposed a technique for distinguishing a few speech codecs. Hic,s'onmez et al. proposed a technique for sound codec distinguishing proof through payload inspecting. They abused the haphazardness and clamorous nature of encoded sound to fabricate factual models that speak to the pressure procedures of various codecs.

They likewise utilized sound quality measures to distinguish pressure follows in sound. P. Bießmann et al. assessed the MP3 PRO parameters. In, MDCT-and MFCC-based highlights are proposed for pressure history investigation. Concerning the genuine situation utilization of pressure history investigation, D. G'artner et al. proposed an effective cross codec confining lattice examination for sound altering recognition, and discharged a database for assessing sound altering location strategies. A prior way to deal with distinguish twofold AMR pressure was proposed in. This investigation utilized measurable highlights including normal sub-band recurrence proportion, normal low recurrence sub-band vitality proportion, bi range highlights and coefficients of a long haul direct indicator. A normal discovery precision of around 87% is accomplished for separating the single and twofold AMR sound. Outlining hand-created highlights is a testing and tedious issue.

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 7, Issue 2, February 2018

III. IMPLEMENTATION STRATEGIES

In this framework, rather than physically extricating the highlights, examination of how to utilize profound learning strategies in this sound legal sciences issue is proposed. Numerous new calculations have been proposed lately in different speech preparing fields, including dialect distinguishing proof, speech acknowledgment, speaker recognizable proof and check. The vast majority of the investigations embrace low-level acoustic highlights (e.g. sub-band vitality or MFCC) as the contribution of the system. Sound imitation more often than not presents just unobtrusive relics (pieces of information), which make the above methodologies imperfect for the current issue. Henceforth, this approach utilizes profound figuring out how to naturally extricate the measurable highlights specifically from the crude sound waveforms.

The proposed work misuses a neural system with dropouts as a classifier to identify sound edges, trailed by larger part voting which helps the execution. This approach accomplished a characterization exactness of 91%. Additionally tries demonstrate that utilizing profound learning for grouping is problematic. This framework, as an augmentation of, researches how to utilize a profound system for programmed include extraction. Another structure in view of Stacked Auto Encoder (SAE) arrange for highlight extraction and a Discrete Wavelet Transform (DWT) for grouping has been proposed. Sound edges are independently utilized as the system input, and the last concealed layer's yield constitutes their highlights. The highlights of the considerable number of edges in the sound clasp are then accumulated and utilized as the contribution of DWT.

This approach enhances the characterization precision to 98%. Not quite the same as conventional encryption-based data security assurance strategies, the rising procedure of data concealing ensures the substance of mystery data as well as cover its reality. Be that as it may, most ebb and flow IH plans are low-limit watermarking calculations in light of static picture or music signals. Significantly less research has been done on high-limit speech incognito correspondence (Steganography). The purposes behind this have two folds: to begin with, the mystery information to be installed (a sentence of speech) are significantly bigger than watermarking circumstance (typically a little twofold picture). Second, human sound-related framework is more touchy than human visual framework, which makes the repetition of cover flag generally considerably littler. With a specific end goal to meet the high-limit prerequisite of speech Steganography, it is fundamental and vital to compress the mystery speech before installing.

Mystery speech is compressed by MELP and by speech acknowledgment. These techniques diminish the mystery bit rate viably, however with a cost of generally high computational multifaceted nature and long time-delay. In this framework, a novel approach is proposed to acknowledge high-limit speech Steganography. To begin with, CS innovation is utilized to compress the mystery speech and lessen its bit rate, which has generally littler calculation stack than condition-of-workmanship mystery speech handling techniques. Mystery bit stream is then implanted into cover speech by SCS quantization methodology. Utilization of Discrete Wavelet Transform (DWT) upgrades both the intangibility and heartiness of the framework.

IV. PROPOSED SYSTEM SUMMARY

The detection of double compressed AMR audio can be regarded as a pattern recognition problem. The traditional framework involves a feature extraction stage and a classification stage. Manual feature extraction relies on known quantization artifacts that discriminate between single and double compressed audio. Compression disturbs various statistics such as band energies and spectrum coefficients, which can be detected by well-designed handcrafted features. However, this process is challenging and time consuming, much effort has been made to allow for automatic learning of features from the data. Here, we adopt a deep learning architecture: the stacked auto encoder.

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 7, Issue 2, February 2018

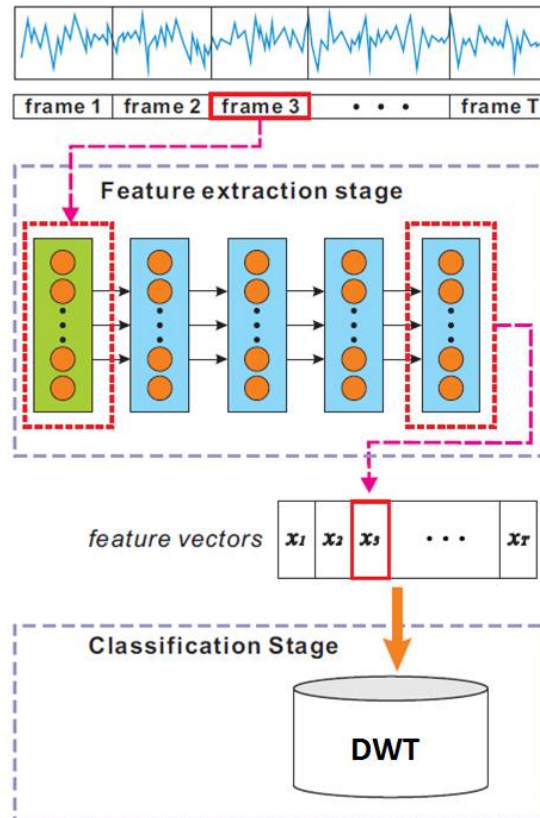


Fig.2 Framework for Identifying Single/Double AMR Audio

Our framework consists of two stages, as shown in above figure. The first stage is feature extraction using an SAE network. The network's input is the normalized waveform samples of an audio frame, and the output of the last hidden layer in the network will be regarded as the feature vector extracted from a given frame. Single and double compressed AMR audio frames are both used to train the SAE network. In the second stage of the framework, DWT is adopted for classification. The inputs of DWT are feature vectors extracted from successive frames from an audio clip. Both single and double compressed AMR clips are used to train DWT.

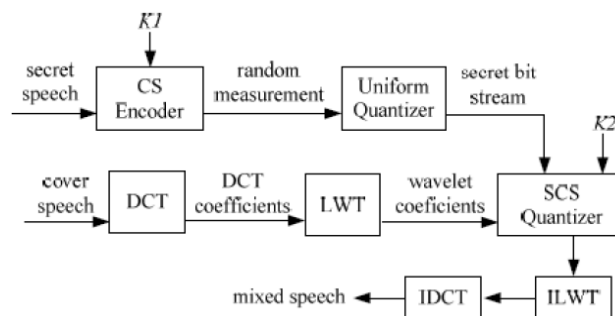


Fig.3 Embedding Framework

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 7, Issue 2, February 2018

In proposed system, the application of Virtual Reality [VR] is utilized, that provides the ability to our framework to generate the key for cryptography, according to that the data is converted to cipher format and the resulting accuracy of virtual reality and its processing is maximum compare to existing strategies.

IV. RELATED STUDY

In the year of 2013, the authors "H.Su, R.Garg, A.Hajj-Ahmad, and M.Wu" proposed a paper titled "Enf analysis on recaptured audio recordings", in that they described such as: Electric Network Frequency (ENF) based forensic analysis is a promising tool for timestamp authentication and forgery detection in multimedia recordings as audios and videos. ENF signal is embedded in an audio recording due to electromagnetic interference from the power lines. The time of creation of a multimedia recording can be determined by comparing the ENF signal embedded in the recording with a reference ENF database collected from the power grid. In their paper, they have conducted a study on effect of recapturing of audio recordings on the ENF embedding, demonstrating that recaptured audio recordings pick up two ENF signals: the content ENF signal which is inherited from the original audio recording; and the recapturing ENF signal which is embedded from the recapturing process.

Conventional ENF signal extraction techniques on such recordings may fail when the two ENF signals are at the same nominal value. A de-correlation algorithm has been used to extract the content ENF signal and the recapturing ENF signal. The experimental results show the effectiveness of the method in the estimation of both the ENF signals.

In the year of 2010, the authors "H.Malik and H.Farid" proposed a paper titled "Audio forensics from acoustic reverberation", in that they described such as: Digital audio recordings can be manipulated by pervasive audio editing software easily. Often forgery would not be naive splicing. Post-processing would be a part of tampering. Post-processing can eliminate the obvious traces of forgery.

Noise can cover audible evidence of forgery and destroy traces of other tampering operations. The detection of additive noise in audio signal is a useful tool for audio forensics. In their paper, they have investigated the effect of additive noise on audio signal, and obtained a feature named "sign change rate" for detecting additive noise. Theoretical analyze and extensive experiments shows the feature is effective in additive noise detection. Also the method can be a potential tool for forgery localization of digital audio. Its only drawback is being very expensive in implementation.

In the year of 2009, the authors "R.Yang, Y.Shi, and J.Huang" proposed a paper titled "Defeating fake-quality MP3", in that they described such as: Compression history identification plays a very important role in digital multimedia forensics. However, most existing literatures mainly focus on digital image forensics, and just a few works consider digital audio. In this paper, they have investigated two popular compression schemes in digital audio, that is, MP3 and WMA, and try to reveal the compression history for a questionable audio signal in the original uncompressed WAV format via analyzing some statistical characteristics of the modified discrete cosine transform coefficients of the audio.

The extensive experimental results have shown that the method can effectively identify whether the given audio has been previously compressed with MP3 and/or WMA, and can further estimate the hidden compression rates, even the compression rate is as high as 128 Kbps (bits per second).

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 7, Issue 2, February 2018

V. EXPERIMENTAL RESULTS

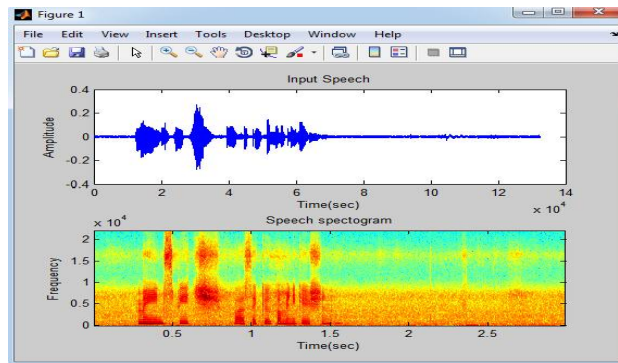


Fig. 4 Input Speech and Speech Spectrogram

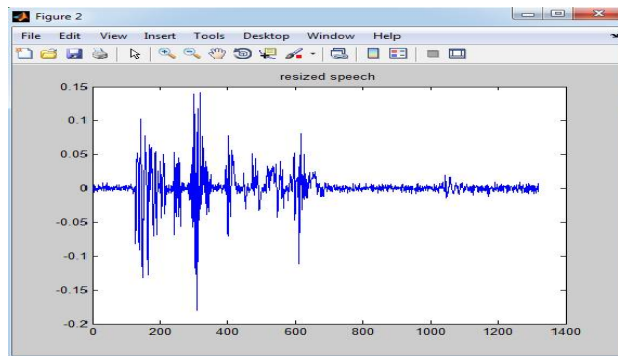


Fig. 5 Resized Speech

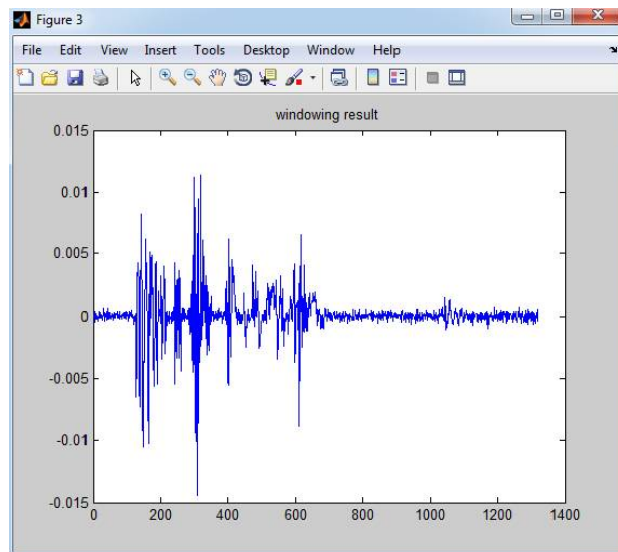


Fig. 6 Windowing

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 7, Issue 2, February 2018

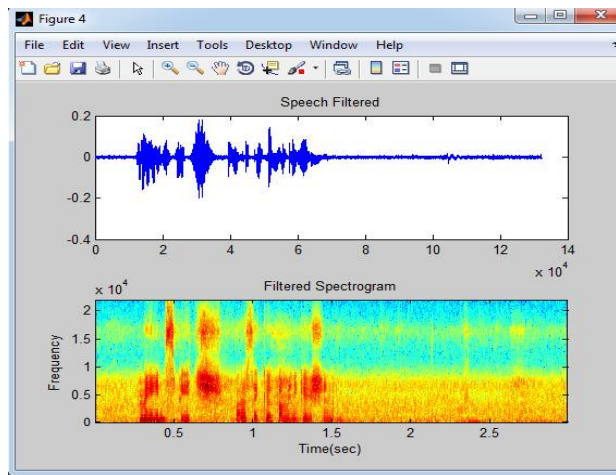


Fig. 7 Filtered Speech and Spectrogram

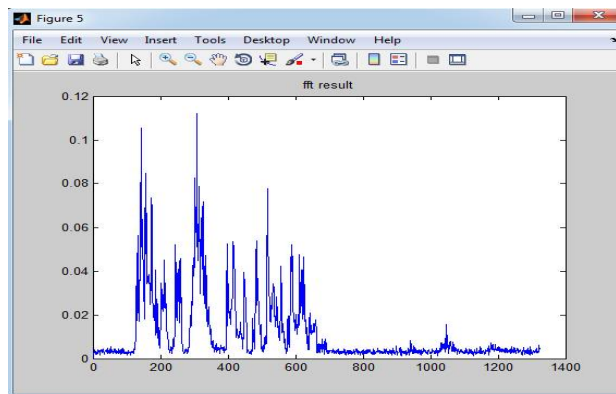


Fig. 8 FFT Result

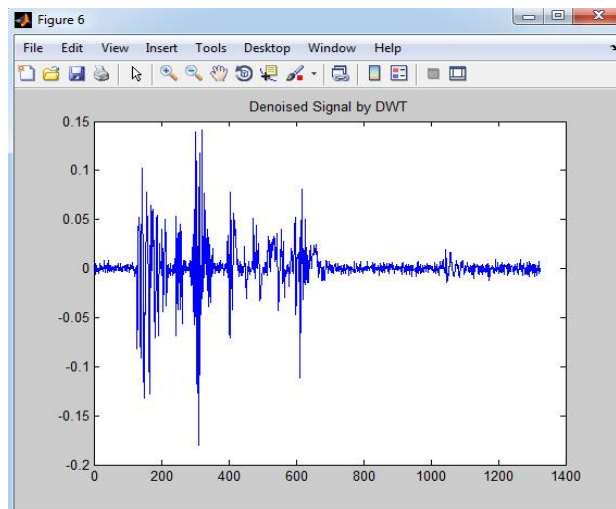


Fig. 9 Denoised Signals by DWT

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 7, Issue 2, February 2018

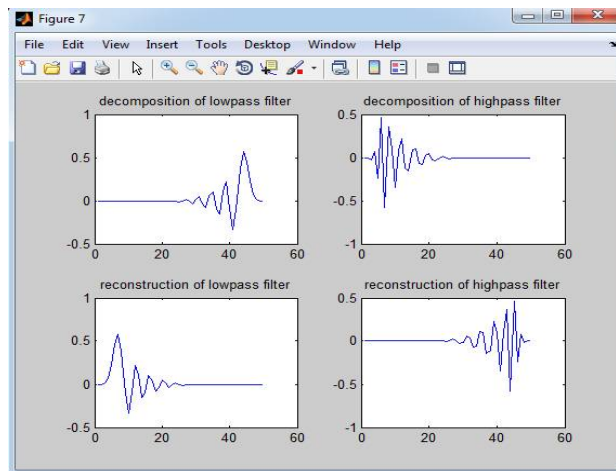


Fig.10 Decomposition and Reconstruction using Lowpass and Highpass Filters

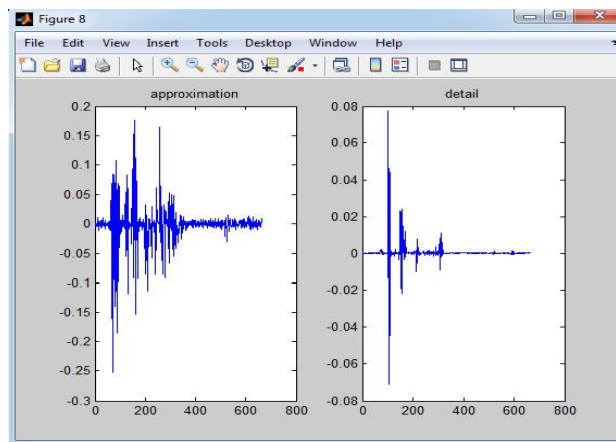


Fig.11 Approximation and 2-Detail

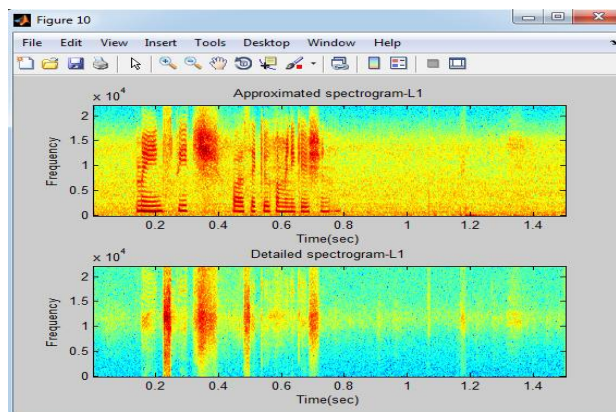


Fig.12 Approximation and Detailed Spectrogram L1

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 7, Issue 2, February 2018

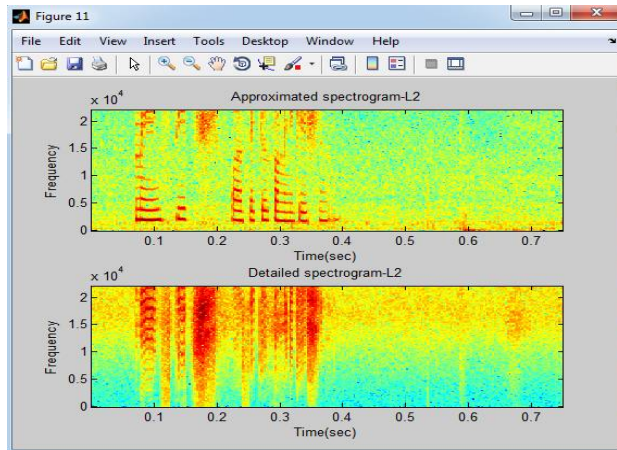


Fig.13 Approximated and Detailed Spectrogram L2

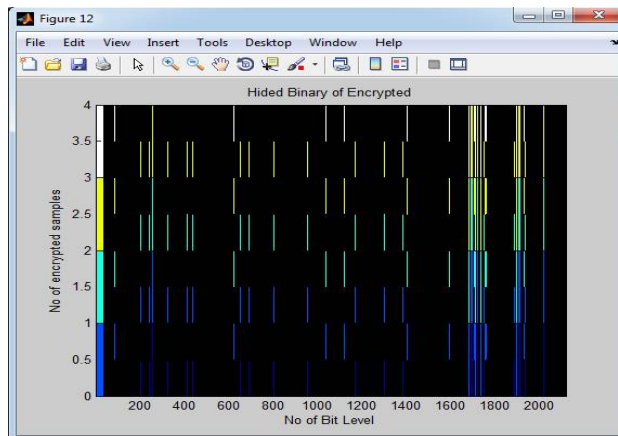


Fig.14 Hidden Binary of Encrypted

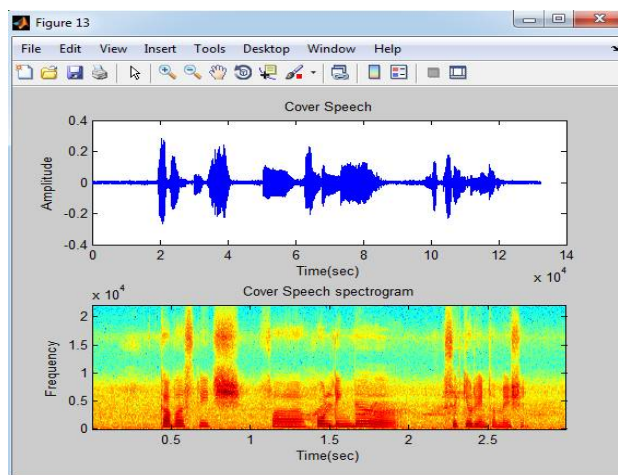


Fig.15 Cover Speech

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 7, Issue 2, February 2018

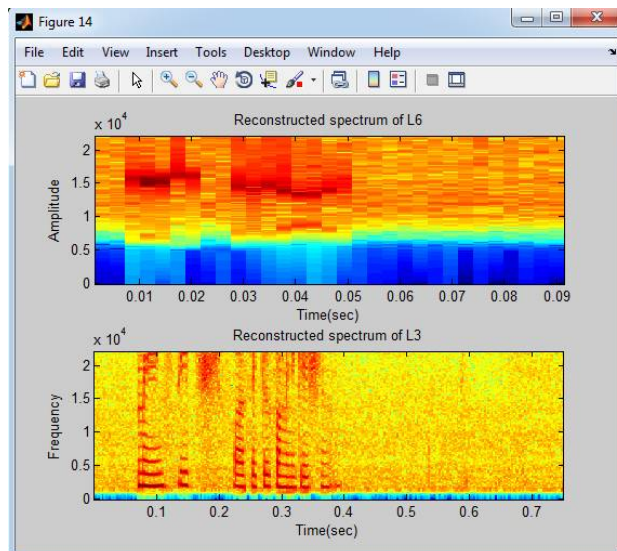


Fig.16 Receiving and Reconstruction

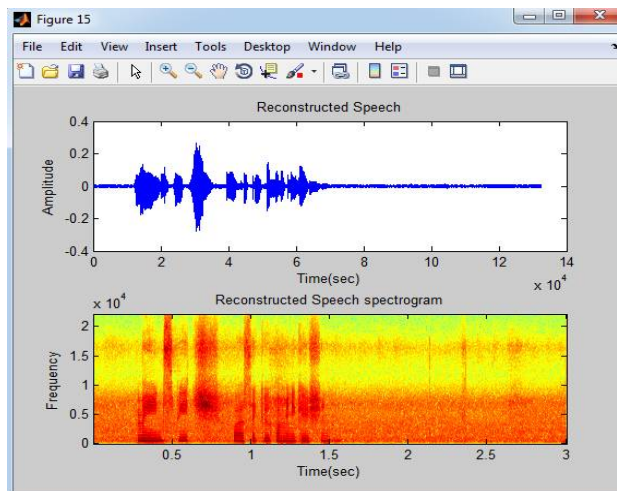


Fig.17 Reconstructed Speech and Spectrogram

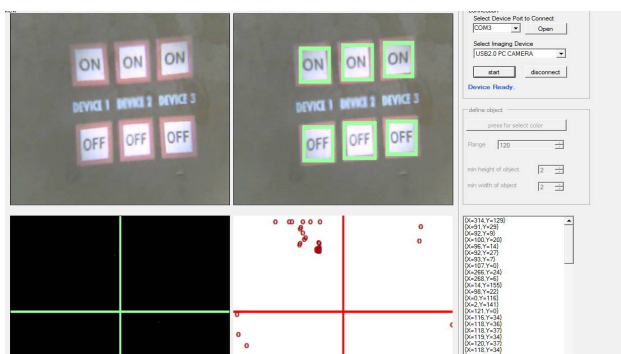


Fig.18 Hardware Prototype – VR Display

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 7, Issue 2, February 2018



Fig.19 Hardware Prototype – VR Setup



Fig.20 Overall System Design - Hardware Prototype

VI. CONCLUSION

The detection of double compressed AMR audio is an important issue in audio forensics. Thus a deep learning technique is proposed to extract the meaningful features from a single audio frame automatically. In this paper, a detection framework based on a carefully designed SAE network combined with DWT is implemented to identify double compressed AMR audio. The features from all the frames in an audio clip are aggregated and classified by DWT and the extracted features from the waveform of audio signal are visualized. The proposed framework is very robust and can be reliably used in various conditions. It is also effective for large databases. The proposed framework

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 7, Issue 2, February 2018

can be achieved a detection accuracy of 98% for 1 second audio clips from the TIMIT database, outperforming existing methods. It can resist frame offset attack and noise contamination.

REFERENCES

- [1] R.Aggarwal, S.Singh, A.K.Roul and N.Khanna (2014), "Cell phone identification using noise estimates from recorded audio," in International Conference on Communication and Signal Processing, pp. 1218–1222.
- [2] P.Bießmann, D.Gärtner, C.Dittmar, P.Aichroth, M.Schnabel, G.Schuller, and R.Geiger (2013), "Estimating MP3PRO Encoder Parameters From Decoded Audio," in GI-Jahrestagung.
- [3] W.Campbell, D.Sturim, and D.Reynolds (2006), "SVM based speaker verification using a gmm supervector kernel and nap variability compensation," in Proc. of the International Conference on Acoustics, Speech and Signal Processing.
- [4] H.Farid (1999), "Detecting digital forgeries using bi-spectral analysis," Report in MIT AI Memo AIM-1657, MIT.
- [5] D.Gärtner, C.Dittmar, P.Aichroth, L.Cuccovillo, S.Mann, and G.Schuller (2014), "Efficient cross-codec framing grid analysis for audio tampering detection," in Audio Engineering Society Convention 136.
- [6] C.Grigoras (2005), "Digital audio recording analysis: The electric network frequency ENF criterion," The International Journal of Speech Language and the Law, vol. 12, no. 1, pp. 63–76.
- [7] H.S.S.Hicsonmez and I.Avcibas (2011), "Audio codec identification through payload sampling," in Proc. of the International Workshop on Information Forensics and Security.
- [8] S.Hicsonmez, E.Uzun, and H.Sencar (2013), "Methods for identifying traces of compression in audio," in Proc. of the 1st International Conference on Communications, Signal Processing, and their Applications, Sharjah, pp. 1–6.
- [9] M.Jahanirad, A.W.A.Wahab, N.B.Anuar, M.Y.I.Idris and M.N.Ayub (2014), "Blind identification of source mobile devices using VoIP calls," in IEEE Region 10 Symposium, pp. 486–491.
- [10] F.Jenner and A.Kwasinski (2012), "Highly accurate non-intrusive speech forensics for codec identifications from observed decoded signals," in Proc. of the International Conference on Acoustics, Speech and Signal Processing, Kyoto, pp. 1737–1740.
- [11] Q.Liu, A.Sung, and M.Qiao (2010), "Detection of double mp3 compression," Cognitive Computation, vol. 2, pp. 291–296.
- [12] D.Luo, W.Luo, R.Yang, and J.Huang (2014), "Identifying compression history of wave audio and its applications," ACM Transactions on Multimedia Computing, Communications, and Applications, vol. 10, no. 3, p. 30.
- [13] H.Malik and H.Farid (2010), "Audio forensics from acoustic reverberation," In Proc. of the International Conference on Acoustics Speech and Signal Processing, pp. 1710–1713.
- [14] M.Qiao, A.Sung, and Q.Liu (2010), "Revealing real quality of double compressed MP3 audio," in Proc. of the international conference on Multimedia, Firenze, Italy, pp. 1011–1014.
- [15] H.Su, R.Garg, A.Hajj-Ahmad, and M.Wu (2013), "Enf analysis on recaptured audio recordings," in Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'13), pp. 3018–3022.
- [16] R.Yang, Z.Qu, and J.Huang (2008), "Detecting digital audio forgeries by checking frame offsets," in Proc. of the ACM Workshop on Multimedia and security, Oxford, United Kingdom, pp. 21–26.