

Scalable Hybrid Speech Codec for Voice over Internet Protocol Applications

Manas Ray

Department of Electronics &
Communication Engineering Birla
Institute of Technology, Mesra
Ranchi, India

Mahesh Chandra

Department of Electronics &
Communication Engineering Birla
Institute of Technology, Mesra
Ranchi, India

B.P. Patil

Department of Electronics &
Telecomm. Engineering, Army
Institute of Technology, Dighi
Pune, India

Abstract—With the advent of various web-based applications and the fourth generation (4G) access technology, there has been an exponential growth in the demand of multimedia service delivery along with speech signals in a voice over internet protocol (VoIP) setup. Need is felt to fine-tune the conventional speech codecs deployed to cater to the modern environment. This fine-tuning can be achieved by further compressing the speech signal and to utilize the available bandwidth to deliver other services. This paper presents a scalable -hybrid model of speech codec using ITU-T G.729 and db10 wavelet. The codec addresses the problem of compression of speech signal in VoIP setup. The performance comparison of the codec with the standard codec has been performed by statistical analysis of subjective, objective and quantifiable parameters of quality desirable from the codec deployed in VoIP platforms.

Keywords—VoIP; Speech Compression; Hybrid Speech Codec; ITU-T G.729 codec; db10 wavelet; Statistical Analysis

I. INTRODUCTION

In the recent years, there has been substantial standardization in the various codec deployment and access technologies for VoIP applications. These standardizations has laid the foundation of the modern 4G technology, which is a packet switched network based on internet protocol and has data rates up to 1 Gbps [1]. This compliments the general characteristics of VoIP setup of ease of launch of additional services along with the conventional telephony –voice services[2].The integration of different web-based applications in modern VoIP service had led to the various challenges of Quality of service(QoS) faced by the VoIP application. These challenges are [3], the requirement of high bandwidth, sensitivity to propagation delay, sensitivity to jitter, etc. Hence, to address these challenges, it is desired to build a highly efficient speech compression technique. The compression is desired to conserve the precious resource of bandwidth, so as to enable the amalgamation of other application in the final VoIP package. The activity of compression of the speech signal is carried out by the codecs deployed in the setup.

Currently, various ITU standard codecs are utilized in VoIP setup for seamless interconnectivity across different systems spanning continents. The data-rates provided by these codecs are the function of engineering adjustments between the quality of voice signals, complexity and bandwidth of the codec [4]. Most commercial VoIP application currently operates on the International Telecommunication Union – Telecommunication

Standardization Sector (ITU-T) standardized codec known as ITU-T G.729 codec based on the principle of Conjugate Structure Algebraic Code Excited Linear Prediction (CS-ACELP)[5].The codec is a hybrid type codec which uses the algebraic sum of the fixed codebook gain vector and the adaptive gain to arrive at the actual codebook vector for performing the linear prediction[6].The codec operates at 8kbps and produces the highest quality of speech reproduction known as “Toll Quality” for most practical and real-time conditions and hence widely deployed in all voice-based operations[6]. This technology provides optimal complexity and is best suited for multimedia digital simultaneous voice and data communication, hence makes it best suitable for VoIP applications[7].The performance of this codec have been extensively tested on various environment, which resulted in concluding that the codec performs consistently with voice signals and narrowband control signals [8], international and regional speech coding standards, channel noise and over degraded transmission channels [9]. However, to accommodate additional services further compression of the speech signal is envisaged.

In the recent years, extensive research has been carried out on the topic of wavelet transforms and its applications in signal processing. The wavelet is defined as a short wave of finite duration, whose average value is zero. It is finite in nature [10].The Wavelet transform represents the signal with very high precision and limited storage requirements [11]. Here the signal is de-composed into component-signals resembling sine-waves, having compressed information in both frequency and time domains. The scaling function $\Phi(z)$ determines the resolution of the analysis and the actual analysis is performed by the mother wavelet function $\Psi(z)$ [12].The definition of the Wavelet transform of the signal $s(z)$ is [13]:

$$W_{\Psi}(p,q) = \int_{-\infty}^{\infty} s(z)\Psi_{pq}(z)dz \\ = \frac{1}{\sqrt{p}} \int_{-\infty}^{\infty} (s(z)\Psi(\frac{z-q}{p}))dz \quad (1)$$

Where

$s(z)$ = original signal

p = scaling parameter

q = translation parameter

The wavelet function is given by

$$\Psi_{p,q} = \frac{1}{\sqrt{p}} \Psi\left(\frac{z-q}{p}\right) \quad (2)$$

The Discrete wavelet transform (DWT) is obtained by sampling the Continuous wavelet transform (CWT) with dyadic grid parameters of translation $q = t$ and the scale $p = 2^j$ and the mother wavelet is defined by

$$\Psi(x) = 2^{j/2} \Psi(2^j x - t) \quad (3)$$

Similarly the scaling function is defined as:

$$\phi(x) = 2^{j/2} \phi(2^j x - t) \quad (4)$$

The original function $s(x)$ can be reconstructed as $s'(x)$ from the scaling and the wavelet functions by the relation [14]:

$$s'(x) = \sum_{t=-\infty}^{\infty} c_t \phi_t(x) + \sum_{t=-\infty}^{\infty} d_{j,t} \Psi_{j,t}(x) \quad (5)$$

Where C_t are the average coefficients and $d_{j,t}$ are detail coefficients. The resultant signal thus generated from the wavelet coefficients as referred in equation 5 above provides a compressed representation of the original signal.

The codec we propose aims to enhance the performance of the CS-ACELP codec. The codec utilizes the concept of CS-ACELP as well as that of the wavelet transform, thereby providing an ideal combination of compression and standardization for deployment in VoIP setup.

The paper is organized into six sections. Section-I introduces the requirement of further compression of the speech signal in modern VoIP applications. Section –II provides a brief background of speech signal compression using wavelets. The concept of the proposed scalable hybrid speech codec is presented in Section-III. Section-IV defines the performance evaluation parameters. The results of the assessments are presented in Section – V. Section VI presents the conclusion.

II. SPEECH SIGNAL COMPRESSION USING WAVELETS

The wavelet transform is a transformation used to study the temporal and spectral properties of non-stationary signals like speech, audio etc. based on the frequency-time multi-resolution property of wavelets [15]. The mother wavelet performs the analysis of the speech signal in wavelet domain. The general criteria for the selection of wavelet family for compression are [16]: a) Availability of minimal support in frequency as well as time domain and b) Availability of a relatively large number of vanishing moments. The greater the number of vanishing moments, quicker is the decay rate of co-efficient leading to compact signal representation [17]. Further, they also provide a higher quality of the reconstructed signal, less distortion and a high degree of compression with a trade-off of higher complexity of operation [18].

The selection process of mother wavelet for the analysis of the signal is governed by the consideration of the quality parameters required in the synthesized signal[19]. In VoIP application, the desirable quality requirements of the speech signal are naturalness of the speech, intelligibility, pleasantness, possibility of recognition of speaker[20]. The desirable parameters of the codecs deployed for such applications include [21-22]: Operation in low bit rate, low complexity, robustness across different speakers and

languages, robustness in performance in the presence of channel errors, etc. The latest research in the field of signal processing has identified that Daubechies family of wavelets provides optimal results in the recognition of spoken digits as compared to other wavelets [15,23,24,25] and has the advantage of approximate shift –invariance and better edge representation as compared to real-valued discrete wavelet transform[26-27]. Hence, the Daubechies family of wavelets has been considered in the proposed codec.

The Daubechies wavelet family is extensively used for analysis of speech signals. They are orthogonal wavelets, with the highest quantity of vanishing moments per given support. Here the scaling and the wavelet functions are not defined[28]. The family of the wavelets is defined as dbN where db stands for family name Daubechies, N represents coefficients generated, similarly the number of vanishing moments being $N/2$ [12,29].

Wavelet transform decomposes the signal into Average coefficients or scaling co-efficient indicating low-frequency components and Detail co-efficients or wavelet –coefficients containing high-frequency components. In the speech signal, high frequencies are present at the onset for a very brief period, while lower frequencies are present later for longer periods [30]. Wavelet transforms resolve all these frequencies simultaneously localized in time to a level proportional to their wavelength, thereby obtaining localization in time as well as frequency. Recent studies of the wavelet decomposition [31] indicates that less than 5% of the maximum value is present in 90% of the wavelet co-efficient and hence can be treated as redundant for the purpose of signal analysis. Compression of the signal is hence achieved by truncating the redundant signals (low valued co-efficient) to zero and reconstructing the signal using the remaining co-efficient [17]. The speech compression technique using wavelet is graphically explained in Fig.1.

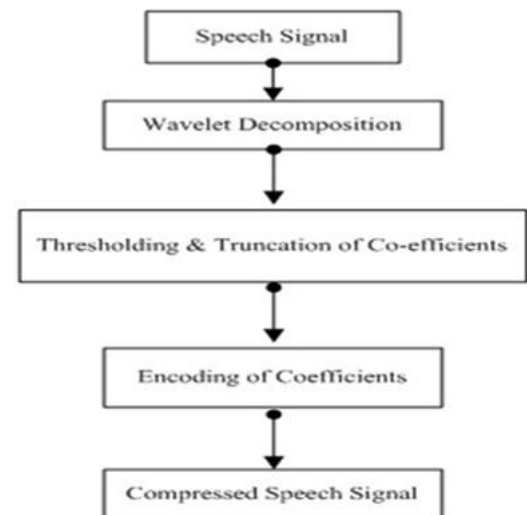


Figure 1. Speech Signal Compression Using Wavelets

The proposed codec utilizes the Daubechies family of the wavelet, specifically db10 wavelet for the wavelet analysis part of the codec.

III. PROPOSED SCALABLE HYBRID CODEC

The work flow diagram of the proposed codec is presented in Fig. 2. In order to compress the bandwidth requirement for the operation of the standard G.729 codec, it is proposed to design a scalable hybrid codec, in which the output of the codec be cascaded to a wavelet-based compression technique to obtain the final compressed reconstructed signal.

The CS-ACELP codec acts as the core-layer for speech signal processing and the wavelet compression acts as the enhancement layer for the required compression. The original signal is fed to a conventional ITU G.729 codec to obtain 8kbps synthesized speech signal. This in-turn is subjected to 5-level recursive wavelet decomposition, since beyond the fifth level of wavelet decomposition, no added advantage is available in for signal processing [11]. We have chosen Daubechies family of wavelets for the decomposition as they concentrate in their approximation coefficients more than 96% of the signal energy [31]. Further, as described in the previous section, Daubechies family of wavelets provides better results in speech recognition, better edge representation for speaker identification and has approximate shift in-variance, which is best suitable for VoIP speech analysis.

The decomposed signal is subjected to thresholding in-order to truncate and de-noise the signal thereby improving the Signal to Noise ratio (SNR) [32]. Adaptive or soft thresholding is used in the proposed codec as this provides the efficient method of de-noising depending upon the signal under study[33]. The thresholding procedure determines the magnitude of compression of the signal[34]. The input signal is approximated from the truncated coefficients by applying inverse wavelet transform. The signal thus synthesized is the compressed version of the original speech signal.

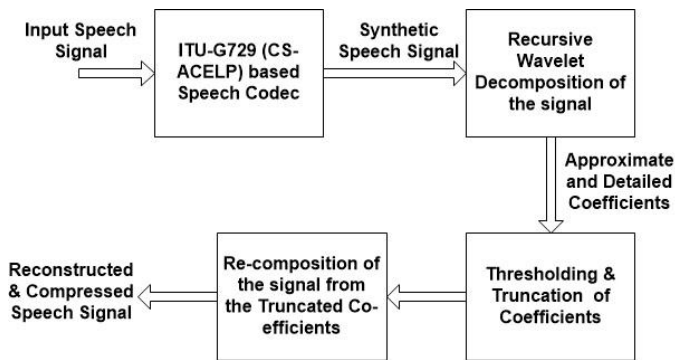


Figure 2. Proposed Speech Codec

As can be observed from the work flow diagram, the complexity of the proposed codec would be bit higher because of the requirement of wavelet transform calculations. The thresholding for compression employed is minimax principle thresholding as available in statistics.

IV. PERFORMANCE EVALUATION PARAMETERS

Subjective and objective test are carried out on a standalone ITU –G.729 CS-ACELP based speech codec and the proposed codec to compare their performance. The test is performed

using 08 samples of speech of both male and female speakers of English and Hindi languages. The age group of the speakers is 30-35 years and is of Indian ethnicity. Calculation of the Mean Opinion Score (MOS) is carried out to determine the subjective evaluation of the speech codec. In MOS test, the original signal and the reconstructed signals are presented to a user, who then provide an acceptability grading between 1 & 5, where 5 is excellent grade [35]. The general rating of CS-ACELP is in the range of 4.1-4.5[36].The objective performance of the codecs is assessed by measuring the parameters of Perceptual Objective Listening Quality Assessment (POLQA), Compression Ratio (CR), Normalized Root Mean Square Error (NRMSE), SNR and the total response time of the algorithm to process the speech signal sample. POLQA [37] is an ITU-T objective speech quality measurement recommendation described as the P.863 recommendation. It analyzes both the original and reconstructed signal sample by sample in the frequency domain with the temporal alignment of the original and the reconstructed signal to provide the MOS mapping [38].The mathematical expressions of the parameters viz. CR, SNR, NRMSE is given below[39-40]:

Compression Ratio is defined as:

$$CR = \frac{\text{Bit length of } (c(k))}{\text{Bit length of } (m(k))} \quad (6)$$

Where, $c(k)$ is the original signal

$m(k)$ is the reconstructed signal

Signal to Noise Ratio is defined as:

$$SNR = 10 \log_{10} \left(\frac{\sigma_x^2}{\sigma_e^2} \right) \quad (7)$$

Where σ_x^2 is the mean square of the original signal

σ_e^2 is the mean square difference between the original and synthetic signal.

Normalized Root Mean Square Error (NRMSE) is defined as:

$$NRMSE = \sqrt{\frac{(c(n)-m(n))^2}{(c(n)-\mu m(n))^2}} \quad (8)$$

Where, $c(n)$ is the original signal,

$m(n)$ is the synthetic signal

$\mu m(n)$ is the mean of the original signal.

The bit rate in kbps is determined as [13]:

$$\text{Bit rate} = \frac{\text{Size of the reconstructed signal file in kilobits}}{\text{Length of the reconstructed signal in seconds}} \quad (9)$$

V. RESULTS

ITU G-729 (CS-ACELP) based codec and the proposed scalable hybrid codec based on CS-ACELP and Daubechies family based wavelet is simulated in MATLAB. The test samples referred in Table-I are iterated against both the codecs and the individual performance is noted.

TABLE I. DETAILS OF SENTENCES USED IN THE EXPERIMENT

Sample No.	Speaker	Language	Sample Sentence
1	Male	English	Welcome to the internet my friend, How can I Help you?
2	Male	English	All systems are running as per predefined parameters
3	Male	Hindi	यहाँ से लगभग पाँच मिल दक्षिण पश्चिम में कटघर गाँव है ।
4	Male	Hindi	धोबिन जब सो कर उठती , तब देखती की चौंका साफ पड़ा हैं और बर्तन मजे हुए हैं ।
5	Female	English	Welcome to the internet my friend, How can I Help you?
6	Female	English	All systems are running as per predefined parameters
7	Female	Hindi	यहाँ से लगभग पाँच मिल दक्षिण पश्चिम में कटघर गाँव है ।
8	Female	Hindi	धोबिन जब सो कर उठती , तब देखती की चौंका साफ पड़ा हैं और बर्तन मजे हुए हैं ।

The quality of the speech codec is evaluated by measuring the MOS, POLQA, Compression Ratio, SNR, and NRMSE for the individual samples. The total response time to process the speech signals is also presented for analysis. These evaluation processes are carried out on Intel i-5 processor with 4GB RAM.

The operational bit-rate observed in the simulation of the proposed codec is 4kbps employing the minimax threshold in MATLAB simulation software. The details of the observation are presented in Table II. The individual results are provided in the following figures. Fig. 3 presents the comparison of the codecs in terms of the MOS, comparison in terms of the Compression Ratio is provided in Fig. 4, Fig. 5 compares the codecs in terms of SNR, analysis of the codecs in terms of NRMSE values are presented in Fig. 6, Fig. 7 compares the codecs in terms of the POLQA and finally the Fig. 8 provides the analysis of the delay response of the codecs. Fig. 9 presents the graphical representation of the original signal and the reconstructed signal output of the proposed codec. The results are summarized in Table III.

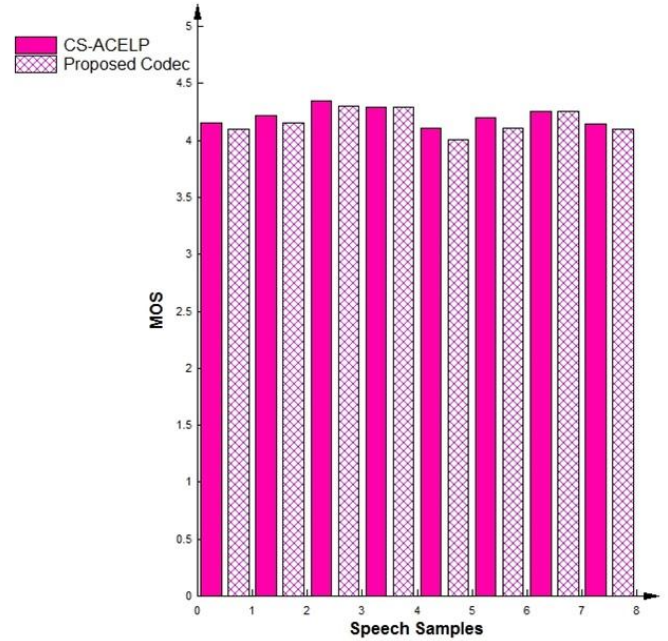


Figure 3. Comparison of MOS

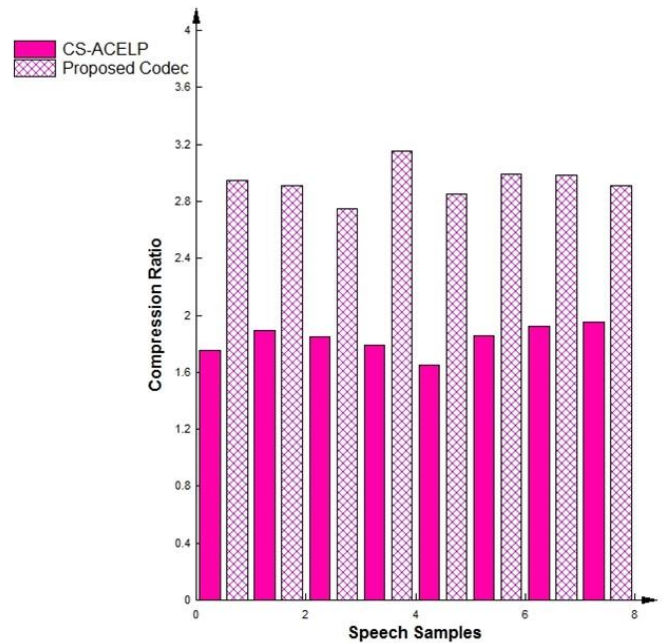


Figure 4. Comparison of Compression Ratio

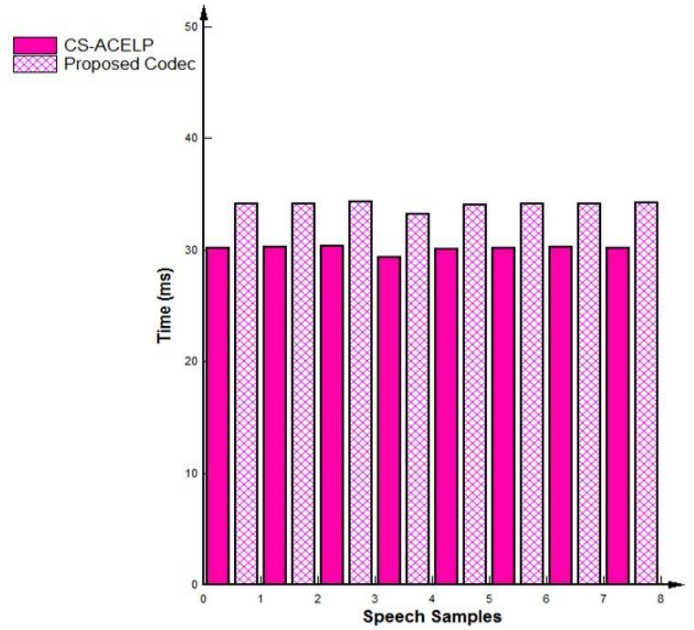
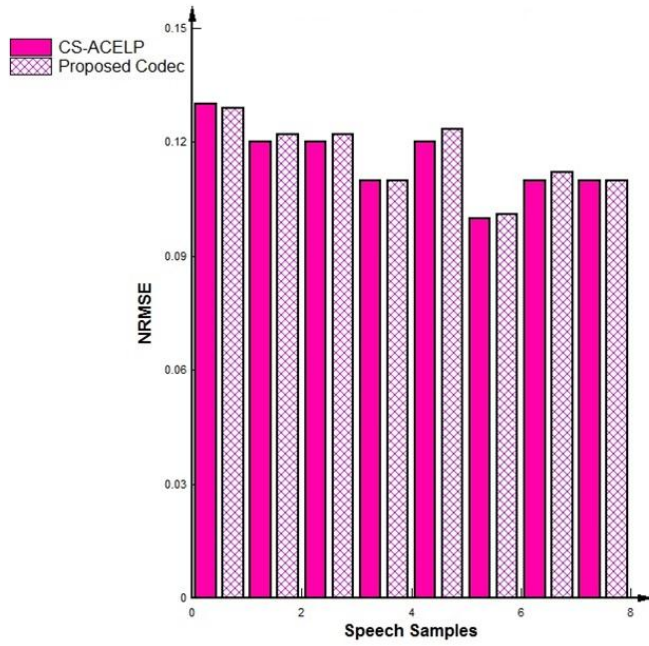
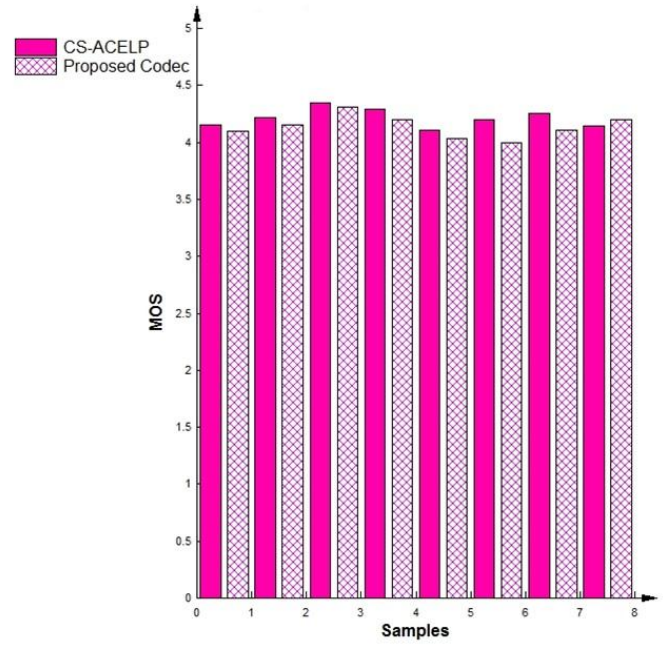
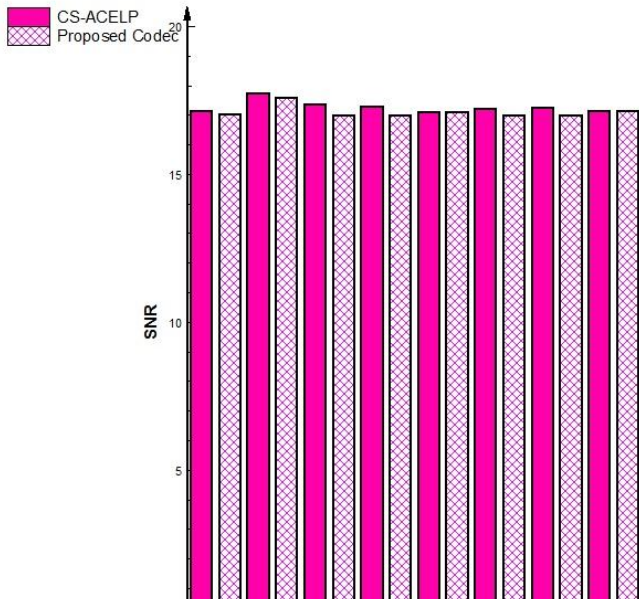


Figure 6. Comparison of NRMSE

Figure 8. Comparison of Delay Response

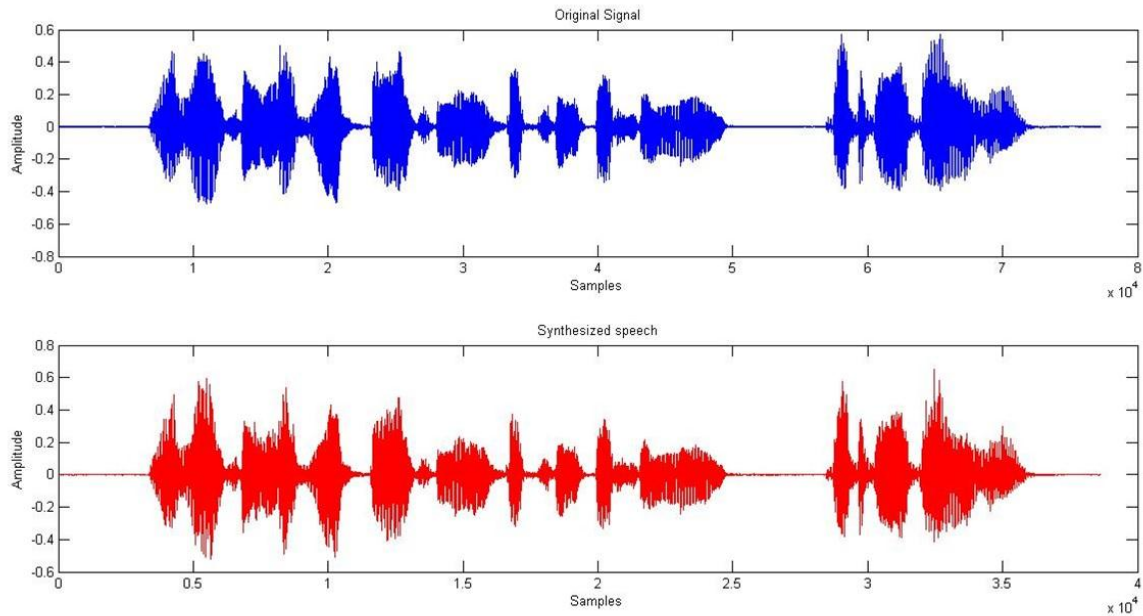


Figure 9. Speech Signal Compression Using Proposed Codec

TABLE II. SUMMARIZATION OF THE BIT RATES OBSERVED USING THE CODECS

Sample	CS-ACELP Codec			Proposed Codec		
	Length of Reconstructed Signal(seconds)	Size of Reconstructed Signal (Kb)	Bit Rate Observed (Kbps)	Length of Reconstructed Signal (seconds)	Size of Reconstructed Signal (Kb)	Bit Rate Observed (Kbps)
1	5	44	8.8	5	18	3.6
2	3	25	8.3	3	11	3.6
3	5	42	8.4	5	17	3.4
4	10	83	8.3	10	41	4.1
5	5	42	8.4	5	19	3.8
6	3	24	8.0	3	11	3.6
7	5	42	8.4	5	19	3.8
8	10	82	8.2	10	41	4.1

TABLE III. SUMMARIZATION OF THE RESULTS OF PERFORMANCE ANALYSIS OF THE PROPOSED CODEC WITH CS-ACELP CODEC

Sr. No	Codec	Bit Rate (Kbps)	Average MOS	Average POLQA	Average SNR (dB)	Average NRMSE	Average CR	Average Delay (ms)
1	CS-ACELP	8	4.33	4.25	17.35	0.13	1.82	30
2	Proposed Codec	4	4.30	4.22	17.30	0.13	2.95	34.75

VI. CONCLUSIONS

The performance of the CS-ACELP based standard codec (ITU-T G.729) is compared with the proposed scalable hybrid codec in this paper. It is observed that proposed codec operates at a bit rate of 4kbps and provides greater degree of compression than the standard codec. The compression can further be fine-tuned to incorporate additional application in the final service delivery by altering the threshold selected for truncating the wavelet co-efficients. It is observed that the proposed codec provides an additional delay to the tune of 5ms which is attributed to higher complexity of calculation of the wavelet transform. Further it is observed that the proposed codec provides comparable results in other parameters under test. It is also observed that the proposed codec provides robust performance across speakers and languages.

The additional delay of 5ms may be treated as acceptable on real-time applications, considering the fact that the codec requires lower bit rate of operation than the ITU-T G.729 standard codec while providing comparable performance in terms of MOS, SNR etc.

Based on the above it can be concluded that the proposed codec provides a viable alternative to the currently deployed codecs in the VoIP setup with an additional feature of high degree of compression as desired in the futuristic VoIP deployments.

REFERENCES

- [1] Reference of 4G[Internet], WikiPedia,2014[updated 2016 February 14; cited 2016 March 10]. Available from: <http://en.wikipedia.org/wiki/4G>
- [2] Sinaepourfard A, Hussain HM, "Comparison of VoIP and PSTN services by statistical analysis". Proceedings of Student Conference on Research and Development(SCOREd).Cyberjaya.pp. 459-461,December - 2011.
- [3] Ray M, Chandra M, Patil BP, " Evaluation of CDMA Microwave Links at Different Environments for VoIP Applications". International Journal of Advance Research in Computer and Communication Engineering (IJARCCCE) . vol. 1, Issue 8, pp. 508-512, October 2012.
- [4] Ray M, Chandra M, Patil BP, "Speech Coding Techniques for VOIP Applications: A Technical Review", World Applied Sciences Journal(WASJ) , vol.33, Issue 5,pp.736-743, May 2015.
- [5] Sat B., Wah B, "Speech and Network -Adaptive Layered G.729 Coder for Loss concealments of Real-Time Voice Over IP", 7th Workshop on Multimedia Signal Processing, pp.1-4, Shanghai, November2005..
- [6] Salami R, Laflamme C, Adoul JP, Kataoka A, " Design and description of CS-ACELP: A toll quality 8kbs speech ceder", IEEE Transctions on Speech and Audio Processing, vol.6, Issue 2, pp.116-130, May 1998.
- [7] Salami R, Laflamme C, Bessette B, Adoul JP. ITU-T G.729Annexure A :Reduced complexity 8kb/s CS-ACELP codec for digital simultaneous voice and data. IEEE Communications Magazine , vol. 35, Issue 9,pp. 56-63, September 1997.
- [8] Parkins ME,Evan K, Pascal D, Thorpe LA. Characterizing the Subjective Performance of ITU-T 8kb/s speech coding algorithm- ITU-T-G.729.IEEE Communications Magazine, IEEE Communications Magazine , vol. 35, Issue 9,pp. 74-81, September 1997.
- [9] Ferraz deCampos Neto S, Karapetian W, "Performance of ITU-T G.729 8kb/s CS-ACELP speech codec with non-voice narrowband signals", IEEE Communications Magazine , vol. 35, Issue 9,pp. 82-91, September 1997.
- [10] Chan YT. Wavelet Basics, Kluwer Academic Publishers,New York,1995.
- [11] Agbinya JI, "Discrete Wavelet Transform Techniques in Speech Processing", IEEE TENCON-Digital Signal Processing Applications,pp.514-519, 1996..

- [12] Somani KP, Ramchandran KI, Resmi NG, Insights into Wavelets from Theory to Practice, Prentice Hill International, New Delhi,2009.
- [13] Kornsing S, Srinonchat J, "Enhancement of Speech Compression Technique Using Modern Wavelet Transforms", IEEE International Symposium on Computer ,Consumer and Control, pp. 393-396,Taichung, June 2012.
- [14] Debdas S, Jagrit V, Chandrakar C, Quereshi MF, "Application of Wavelet Transform for Speech Processing", International Journal of Engineering Science and Technology(IJEST), vol.3, Issue 8,pp.6666-6670, August 2011.
- [15] Elrgaby M, Amoura A, Ganoun A, "Spoken Arabic Digits Recognition Using Discrete Wavelet.", IEEE 16th International Conference on Computer Modelling and Simulation(UKSIM),pp. 275-279,Cambridge, March 2014.
- [16] Vishwanath V, Anderson W, Rowlands J, Ali M, Tewfik A, " Real-time Implementation of a Wavelet Based Audio Coder on the T1 TMS320C31 DSP Chip",5th International Conference on Signal Processing Applications & Technology (ICSPAT), TX.,October 1994
- [17] Strang G, Nguyen T, Wavelets and Filter Banks, Wesley-Cambridge Press ,USA,1996.
- [18] Elaydi H, Jaber MI, Tanbours MB, "Speech Compression Using Wavelets", [Internet]. 2010[cited January 2016] Available from : <http://site.iugaza.edu.ps/helaydi/files/2010/02/Elaydi.pdf>
- [19] Fahmy MF, El-Raheem GMA, Yaseen YM, Sommia AA., "On the Optimal Choice of wavelet bases and Signal Compression Using the wavelet lifting Scheme",IEEE Proceedings of the Nineteenth National Radio Science Conference (NRSC2002) , pp.434-441, 2002.
- [20] Goode B, "Voice over Internet Protocol (VoIP)", Proceedings of IEEE, vol. 90, Issue 9, pp. 1510-1517, September 2002.
- [21] Delprat M, Urie A, Envi C, "Speech Coding Requirements From The Prespective of The Future Mobile Systems",Proceedings of IEEE Workshop on Speech Coding for Telecommunications,pp. 89-90, 1993.
- [22] Benesty J.Springer , Handbook of Speech Processing, Springer-Verlag , Berlin,2008.
- [23] Karam JR, Philips WJ, Robertson W, "Optimal feature Vector for Speech Recognition of Unequally Segmented Spoken Digits", IEEE Canadian Conference on Electrical & Computer Engineering, pp.327-330,Halifax, 2000.
- [24] Krishnan VRV, Jayakumar A, Anto PB, "Speech Recognition of Isolated Malayalam Words Using Wavelet features and Artificial Neural Networks", 4th IEEE International Symposium on Electronic Design Testing and Applications, pp. 240-243, Hongkong, 2008.
- [25] Joseph SM, "Spoken Digit Compression Using Wavelet", IEEE International Conference on Signal and Image Processing, pp. 255-259,Chennai, December 2010.
- [26] Khare M, Prakash O, Srivastava RK, Khare A, "Daubechies Complex Wavelet Transform based approach for Multiclass Object Classification",IEEE International Conference on Control, Automation and Information Sciences (ICCAIS); pp. 206-211,Gwanju, December 2014.
- [27] Noureddine A, Bousselmi S, Cherif A, "Optimized Speech Compression Algorithm based on Wavelets Techniques in Realtime Implementation on DSP", International Journal of Information Technology & Computer Sciences (IJITCS),vol.7, Issue 3, pp. 33-41, February 2015.
- [28] Reference on Daubechies Wavelet[Internet]. Wikipedia;2009 [updated 2016 March 09; cited 2016 March 10].Available from: http://en.wikipedia.org/wiki/Daubechies_wavelet
- [29] Daubechies I, "Ortho-normal bases of compactly supported wavelets", IEEE Communication on Pure and Applied Mathematics, Volume 41, pp. 906-996, 1998.
- [30] Rabiner L, Juang B,Fundamentals of Speech Recognition, Prentice Hall International Inc.,New Jersey,1993.
- [31] Kinsner W, Langi A, "Speech and Image Signal Compression using Wavelets",IEEE Wescanex Conference Proceedings,pp. 368-375, Saskatoon, 1993.
- [32] Sunny S, Peter SD, Jacob KP, "A New Algorithm for Adaptive Smoothing of Signals in Speech Enhancement". In Lee Garry Editor.

- International Conference on Electronic Engineering and Computer Science., pp.339-343, Beijing: Elsevier Science Direct,May 2013.
- [33] Ruwei L, Changchun B, Bingyin X, Moashen J., “Speech Enhancement Using the Combination of Adaptive Wavelet Threshold & Spectral Substraction Based on Wavelet Packet Decomposition” IEEE 11th International Conference on Signal Processing (ICSP),pp. 481-484, ,Beijing, 2012.
- [34] Farouk MH. Application of wavelets in Speech Processing, Springer Science & Business Media, USA, 2014
- [35] Ray AK, Acharya T. Information Technology: Principles and Applications, Prentice-Hill of India Private Limited, New Delhi ,2004.
- [36] Nagireddi S, VOIP Voice and FAX Signal Processing, John Willey& Sons Inc. New Jersey 2008.
- [37] Recommendation P.863: Perceptual Objective Listening Quality Assesment[Internet].ITU-2009[updated 2016 March 04;cited 2016 March 10].Available from: <http://www.itu.int/rec/T-REC-P.863/en>
- [38] Reference on Perceptual Objective Listening Quality [Internet].Wikipedia;2014[updated 2015 October 26;cited 2016 February 24]. Available from: <https://en.wikipedia.org/wiki/POLQA>
- [39] Ambika D, Radha V, “A Comparative Study between Discrete Wavelet Transform and Linear Predictive Coding”, IEEE World Congress on Information and Communication Technologies, pp.965-969,Trivandam,November 2012.
- [40] Najih AMMA, Ramili AR, Ibrahim A, Sayed AR, “Comparing Speech Compression Using Wavelets with other Speech Compression Schemes”. IEEE Proceedings of Student Conference on Research and Development. pp. 55-58., 2003