Efficient Signal Quality Improvement Approach for Wideband VoIP System in Pervasive Environment

[1]H P SINGH, [2]SARABJEET SINGH, [3]JASVIR SINGH
[1,2]Department of Physics, Dr B R Ambedkar National Institute of Technology, Jalandhar, India
[3]Department of Electronics Technology, Guru Nanak Dev University, Amritsar, India
harjit_nit@yahoo.co.in

Abstract: -Voice over Internet Protocol (VoIP) is a the best alternative to the traditional voice communication system, since VoIP system converges the data, voice and video data and reduces the cost of call transmission by passing packets through the available bandwidth through internet protocol. Since the environmental noise may affect the intelligibility of the speech signal, so it is desirable to use the speech processing methods to enhance the quality and intelligibility of the speech signal. The proposed scheme, interpolated finite impulse response (IFIR), is implemented as post-processor after decoding the signal in wideband VoIP system. The performance of the proposed scheme is evaluated for various types of noises at different network conditions. The results of the proposed scheme are measured with the wideband-extension to perceptual evaluation of speech quality measurement (WB-PESQ) for wideband signal. The performance of the proposed system is compared with the existing techniques for quality improvement in VoIP system. The results show much improvement in speech quality with proposed scheme in comparison to other similar schemes.

Key-Words: Wideband VoIP, AMR-WB, IFIR, Background Noise

1 Introduction

Voice over Internet Protocol is a popular communication service for transporting voice data packets over packet switched networks such as internet. VoIP reduces the cost of call transmission by passing voice and video packets through the available bandwidth for data packets and also provide additional services such as voice and video conferencing, text chat, caller ID, voice mail, call forwarding [1]. But the quality of speech signal is one of the main problems in the implementation of the voice over IP. The speech quality is not only affected by IP impairment such as packet loss etc, but also due to the presence of background noise since the VoIP user may be in wired or in mobile environment. The acoustic noise affecting the voice signal quality may include variety of different components. This type of noise is typically non-stationary, may have an average spectrum close to the user's own voice [2]. Since the environmental noise may affect the intelligibility of the speech signal, so it is desirable to use the speech processing methods to enhance the signal quality and intelligibility. However, to be able to compete with the highly reputed public switched telephone network (PSTN), VoIP system should be able to achieve comparable quality to that achieved by PSTN. Many approaches to improve the speech signal and optimization of the codec can be found in the literature. The spectral subtraction based noise suppression system was proposed by Boll [3]. The Kalman filtering method was proposed for speech enhancement in [4] and the performance of the proposed method was compared with stationary and non-stationary Wiener filtering method. Langi [5] had implemented ITU-T G.723.1 voice codec algorithm for VoIP gateways on TMS320C5402 DSP processor and use optimization techniques to improve processing time. Han et.al [6] had implemented the Weiner filter operation for noise reduction in VoIP speech codecs. A modified wiener filter based noise reduction scheme optimized to the estimated SNR at each frequency bin as a logistic function. Malvar et.al [7] presented the use of optimal FIR pre and post filters for decimation and interpolation of the random signals. Wu et.al [8] proposed the use of the multipath adaptive interpolated FIR filter for the echo cancellation for the communication systems. The finite impulse response filters have been investigated and implemented on TMS320C6713 DSP processor for quality improvement of the degraded VoIP signal [9] . The interpolation finite impulse filter for noise reduction and concealment of the lost packets was also analyzed and implemented on TMS320C6713 DSP processor for VoIP system and the implementation results indicates much improvement in the signal quality of the VoIP signal [10, 11]. This paper describes the study and performance evaluation of Interpolated finite impulse response (IFIR) filter based system for wideband VoIP applications in the presence of the background noise.

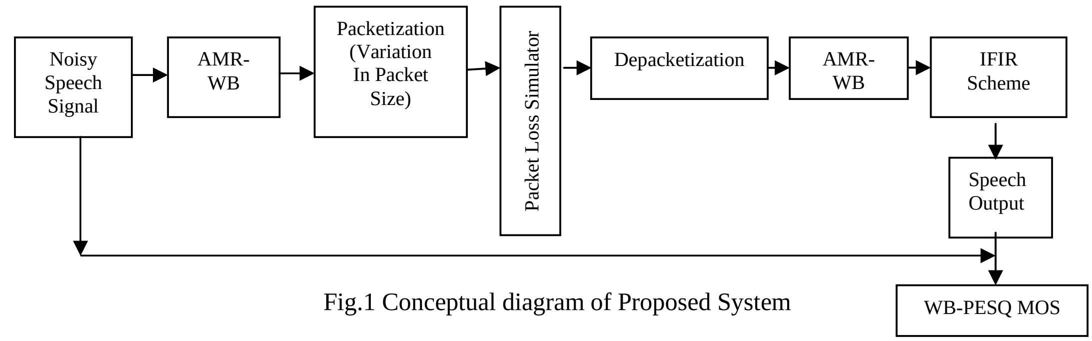


Fig.1 Conceptual diagram of Proposed System

The noisy speech signal is encoded with Adaptive multi-rate wideband (AMR-WB) [12] codec and the efficiency of the proposed system is tested at varying packet size & varying packet loss rate network conditions. The results of the proposed system are then compared with the system which involves interleaving technique.

The brief description of the proposed System, filter design is presented in the next section. The packet loss modeling is described in Section III and the Section IV presents the design of the IFIR scheme. The performance analysis results of the proposed scheme are presented in Section V. The last section concludes the work and presents the future work.

2 Proposed System

To conceal the lost packet and reduce the effect of environmental noise, the IFIR based scheme is proposed. The IFIR filter is applied as post processor after the decoding in the proposed system. The performance of the proposed system is evaluated with Wideband extension- Perceptual Evaluation of Speech Quality (WB-PESQ) measurement defined by ITU-T recommendation P.862.2 [13] for wideband coders. The proposed system is presented in Fig. 1. The simulation for VoIP system is performed in noisy environment with different types of background noises. The basic steps in derivation of the proposed system are:

Step 1: The noisy speech signal is fed in to the system, which is degraded with various types of background noises including babble, car and street at SNR of 0, 5, 10, 15 dB. The about 500 noisy and clean speech samples at different SNRs for both male and female are taken from [14].

Step2: The noisy speech signal is then encoded with AMR-WB speech encoder at different data rates, which is the compressed version of the input signal.

Step3: To save the bandwidth and to reduce the header overhead in VoIP system, the multiple speech frames are transmitted in single packet [15]. In case of the larger size, there is a distortion in the voice quality level [16]. The compressed signal is then packetized into VoIP packets to transfer it to the IP network. To check the network efficiency, the packet size is varied from one to six voice fames into single packet.

Step4: To introduce the IP network impaired into speech signal, the Gilbert model [17]-[19] is used in this work. The signal is degraded with different packet loss rates.

Step5: The degraded VoIP signal is depacketized and then decoded with AMR-WB decoder.

Step6: The proposed IFIR scheme is implemented on the degraded VoIP signal to enhance the signal quality.

Step7: The signal measurement is performed with WB-PESQ. After comparing the degraded signal with the original, the WB-PESQ measurement gives the subjective measurements.

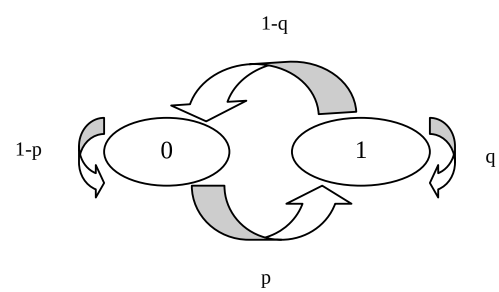


Fig.2 Two-state Gilbert Model

3 Packet Loss Modeling

Packet loss is a major source of speech impairment in voice over IP networks. Packet losses are not independent on a frame-by-frame basis, but appear in bursts. Such a loss may be caused by discarding packets in the IP networks (network loss) or by dropping packets at the gateway/terminal due to late arrival (late loss). Network loss is normally caused by congestion (router buffer overflow), routing instability such as route changes, link failure, and lossy links such as telephone modems and wireless links. Congestion is the most common cause of loss.

The packet loss behavior of IP networks can be represented as a Markov process because several of the mechanisms that contribute to loss are transient in nature (e.g. network congestion, late arrival of packets at a gateway/terminal, buffer overflow or transmission errors) [20]. Several models [21, 22] have been proposed for modeling network loss characteristics. A discrete Markov chain with a set of M states $S = (S_1, S_2, \dots, S_M)$ characterizes the course of the process with regard to the current state, which may change over time at predefined events, such as packet arrivals, based on transition probabilities. Each state is associated with different error or packet loss behavior. Let q_t denote the current state at event time $t, t \in N_0$. Then the probabilities a_{ij} to change from state $q_{t-1} = i$ to $q_t = j$ are given in the transition matrix A

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1M} \\ \vdots & \ddots & \vdots \\ a_{M1} & \cdots & a_{MM} \end{pmatrix} \quad (1)$$

With coefficients

$a_{ij} = P\left(q_t = S_j \mid q_{t-1} = S_i\right), \; i \le M, \; j \le M,$

Where $\sum_{j=1}^{N} a_{ij} = 1; \qquad a_{ij} \ge 0$

The steady states probabilities can be found as:

$$\pi_k = \sum_{j=1}^{M} \pi_j a_{jk}, \;\; k = 1, \dots, M; \qquad \sum_{j=1}^{M} \pi_j = 1 \quad (2)$$

The error or packet loss rates in each state $E = \left(e_1, \dots, e_M\right); 0 \le e_j \le 1$ and the output of the process $O\left(t\right)$ as a binary sequence $O\left(t\right) \in 0,1$ indicating an error or loss at an event with $O\left(t\right) = 1$, whereas $O\left(t\right) = 0$ stands for error free events, respectively.

Most research in VoIP network uses a Gilbert model to represent packet loss characteristics [17]-[19]. In 2-state Gilbert model as shown in Fig.2, there are two states (state *0* and state *1*). The state *0* represents that a packet being correctly received and state *1* represents that a packet being lost. Let p be the transition probability for the network model to drop a packet given that the previous packet is delivered i.e. the probability for network model to go from state *0* to state *1*. Let q is the probability for the network model to drop a packet given that the previous packet is dropped, i.e. the probability for the network model to stay in state 1. This probability is also known as the conditional loss probability. Let $p0$ and $p1$ denote the probability of the network model to be in state *0* and *1*. The probability for a packet to be dropped regardless whether the previous packet is delivered or dropped i.e. the unconditional loss probability is exactly the probability for the network model to be in state *1* $(p1)$,

$$p0 = \frac{q}{p+q}; \qquad p1 = \frac{p}{p+q} \quad (3)$$

The transition matrix is given as

$$P = \begin{pmatrix} 1-p & p \\ q & 1-q \end{pmatrix} \quad (4)$$

The speech signal of the VoIP system is degraded at different packet loss rates at 2%, 4%, 8% and 10% as shown in theTable I.

Table I Simulated Loss Rates

PLR (%)	p	q
2	0.0032	0.15
4	0.012	0.25
8	0.025	0.25
10	0.10	0.85

4 Interpolated FIR Filter Design

The multirate filtering is a novel method to save number of arithmetic operations for FIR filter designs. The filter is implemented as a cascade of the two FIR sections, where one section generates the sparse set of impulse response values with every L^{th} sample being non-zero and other section performs the interpolation. The interpolated finite impulse response (IFIR) filter requires approximately $(1/L)^{th}$ of the multipliers required for conventional equivalent FIR filter. Multirate filtering is effective in narrowband and wideband applications. Since the internal data rate in IFIR filters is constant, so there is no problem of internal aliasing which is one of the major design considerations in multirate filtering [23]. The model filter $H_m(z)$ with impulse response $h_m(n)$ is considered [24, 25]. The $(L-1)$ zero-valued samples are inserted between the original samples of $h_m(n)$. The up-sampled sequence $h_m^{'}(n)$ is:

$$h_m^{'}(n) = \begin{cases} h_m(n/L), & n = iL, i = 0, \pm 1, \pm 2, \dots \\ 0, & otherwise \end{cases} \quad (5)$$

The Z-transform of $h_m^{'}(n)$ is:

$$H_m^{'}(z) = H_m(z^L) \tag{6}$$

The implementation of $H_m^{'}(z)$ is obtained from the implementation of $H_m(z)$ by replacing each delay with L delays. The interpolated impulse response is generated by cascading $H_m(z^L)$ with an interpolator $G(z)$. The block diagram of the IFIR filter is presented in Fig.3. The overall frequency response becomes:

$$H_i(z) = H_m(z^L)G(z) \tag{7}$$

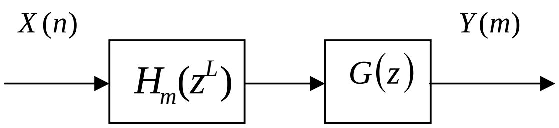


Fig.3 Interpolated Finite Impulse Response Filter

(a) Expansion Factor (L)

The expansion factor L must be a positive integer value of which must be equal or greater than 2 and less than maximum value of $L_{\max}$,

$$2 \le L \le L_{\max} \tag{8}$$

The maximum possible expansion factor $L_{\max}$ is determined using the following relationship:

$$L_{\max} = [\pi / \omega_{SL}] \tag{9}$$

where, ω_{SL} is the stop band edge frequency of the low pass IFIR filter. It is recommended to select L, somewhat smaller than $L_{\max}$ to avoid the more complex structure of the image filter [24, 26]. As the expansion factor increases, the order of the model filter decreases. Consequently, a higher order image filter is needed to remove replicas.

(b) Image Filter $G(z)$

The purpose of the image filter $G(z)$ from time-domain point of view is exactly that computing what the zero-valued samples of $H_m(z^L)$ should be. From frequency-domain point of view, interpolator $G(z)$ must attenuate the replicas of $H_m(z^L)$ in frequency domain. From both points of view it is advantageous to have the order of $G(z)$ as small as possible, in order to achieve the computational efficiency of IFIR implementation [26]. To reduce the order of $G(z)$ to the smallest possible, it is important to carefully consider band edge frequencies. The $G(z)$ filter must pass everything below the f_{pass} frequency:

$$f_{pass,\text{int}} = f_{pass} \tag{10}$$

The image filter $G(z)$ must suppress the closest replica of $H_m(z^L)$ starts at $(1/L) - f_{stop}$.,therefore:

$$f_{stop,\text{int}} = \left(\frac{1}{L}\right) - f_{stop} \tag{11}$$

(c) Design Steps:

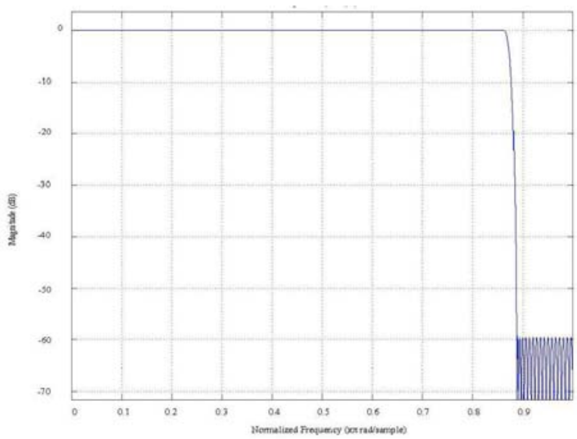
The process of the designing an IFIR filter summarized as:

a) Select the suitable expansion factor L.
b) Design the model filter $H_m(z)$
c) Up-sample the model filter $H_m(z)$ by L to create $H(z^L)$
d) Design the image filter $G(z)$ to remove replicas.

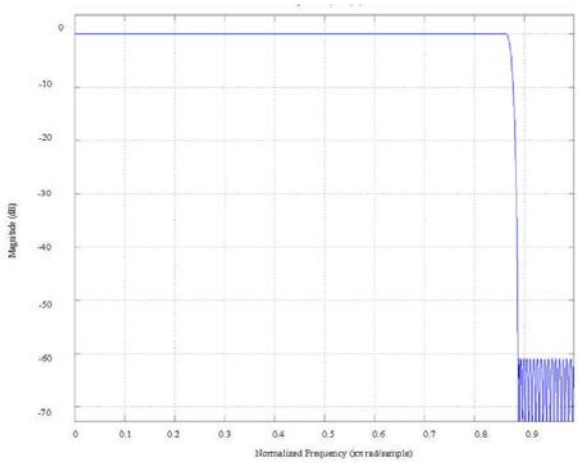
(d) Computational Complexity:

The computational complexity is computed in terms of the number of multipliers needed for implementation of the digital filter [24, 25]. In IFIR filter the passband and stopband widths are only $(1/L)^{th}$ of those of the model filter. The effect of the interpolation of the impulse response is to shrink the passband and transition bands without any significant increase in the number of arithmetic operations. The length of the required FIR filter for given specifications is approximately [27]:

Where, δ_1 and δ_2 are the passband and stop ripples and ω_p and ω_s are the passband and stopband edge frequencies. The proposed IFIR scheme for VoIP speech signal improvement is designed using the MATLAB. For wideband system, the cutoff frequency is 6900 and sampling frequency is16000 Hz. The number of multipliers required are 148 with expansion factor L=2, in comparison to the conventional FIR filter which require 262 number of multipliers and the computational complexity is reduced by 45% with IFIR filter in comparison to the conventional filter. The passband and stopband ripples for the designed filter are 0.001 and 0.001 resp. The Fig.4 and Fig.5 present the magnitude response and impulse response of the designed filter designed with conventional design technique and with IFIR design technique. This computationally efficient IFIR filter is suitable for the VoIP applications and it significantly reduce the noise content in the speech signal.

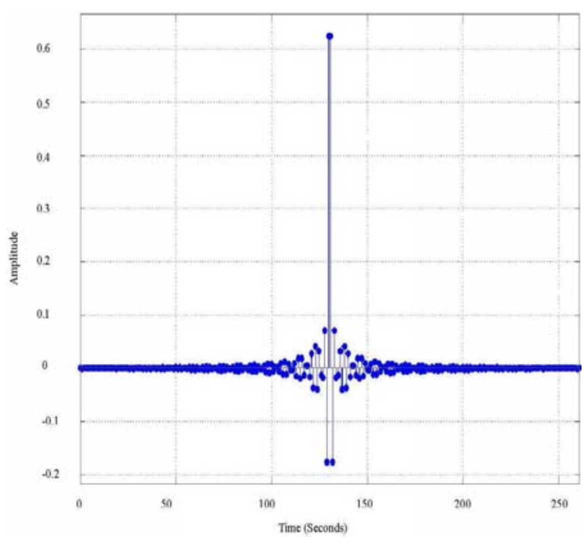


(a)

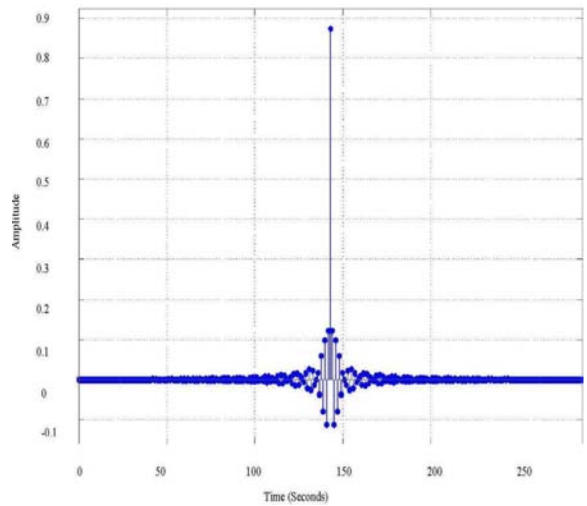


(b)

Fig 4. Frequency responses (a) FIR filter $H(z)$ (b) IFIR filter $H_i(z)$



(a)



(b)

Fig 5. Impulse responses (a) FIR filter $H(z)$ (b) IFIR filter $H_i(z)$

5 Results

The performance results of the proposed system are evaluated for varying packet loss rates and for VoIP packet sizes in various noisy conditions as discussed here below:

(I) Variation of Packet size & Packet loss rate

The performance of the proposed IFIR based noise reduction system is evaluated for wideband VoIP system at varying packet sizes and varying packet loss rates. The packet size used in during the simulation varied from one voice frame per packet to six voice frames per packet and the signal is degraded with packet loss at various packet loss rates varying from 2% to 10%. The average gain in WB-PESQ MOS scores at each packet size for different packet loss rates is plotted in Fig.6 and Fig.7 for AMR-WB 15.85 kbps and AMR-WB 18.25 kbps respectively. The proposed system is effective not only for single voice frame in each packet transmission but also very much effective for the multiple voice frames in each packet. The multiple frame transmission leads to the saving of bandwidth in the network which can be used for other purposes in the VoIP system. The significant increment in MOS scores is achieved for higher packet loss rate, as presented in Fig.6 and Fig.7.

(II) Evaluation for various noise types

The performance of the proposed system is evaluated for various types of the noisy conditions including babble, car and street noise at different SNR 0, 5, 10 and 15 dB in wideband VoIP system. At various packet sizes, the gain in the WB-PESQ

MOS scores for each noise type is plotted against varying packet loss rates in Fig. 8 and Fig. 9 for AMR-WB 15.85 kbps and AMR-WB 18.25 kbps respectively. In case of babble noise, AMR-WB at 15.85 kbps gives better results at high SNR with all tested packet sizes. In case of car noise for small packet sizes gives good results at low SNR and for large packet size, the high SNR gives significant increment in MOS scores. Low value of SNR is better in case of street noise with AMR-WB 15.85 kbps as given in Fig.8. For AMR-WB 18.25 kbps, the increasing value of SNR is preferred with increasing packet sizes with babble noise type. The much better results can be found at low SNR values at all tested packet sizes with car and street noises. The results of AMR-WB 18.25 kbps are presented in Fig.9.

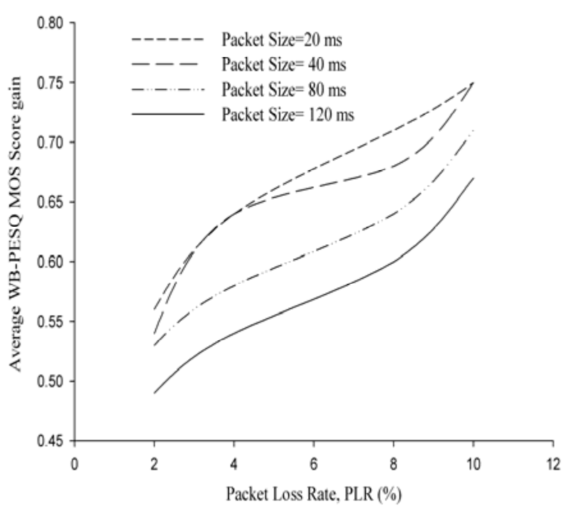


Fig.6 Signal coded with AMR-WB 15.85 kbps

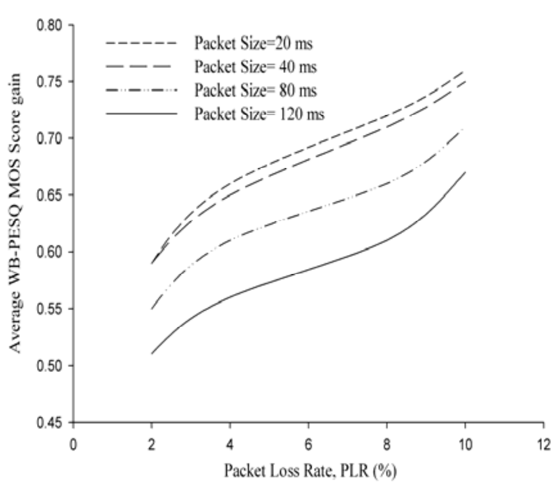
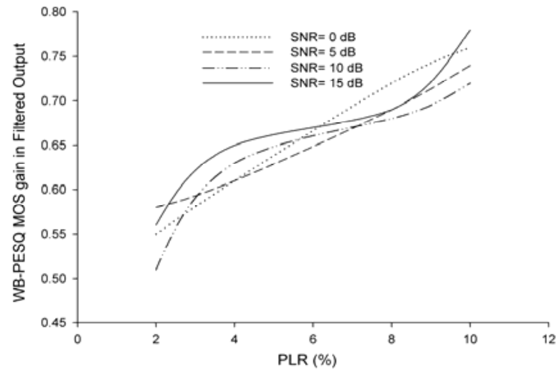
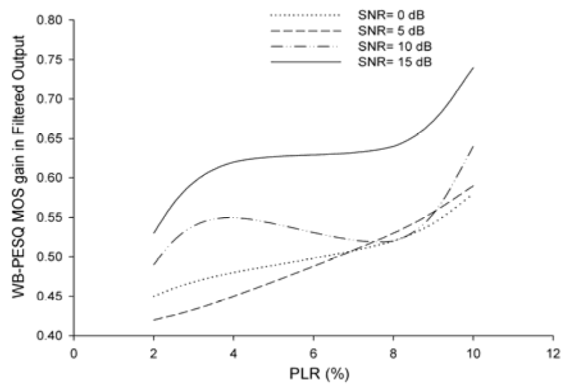


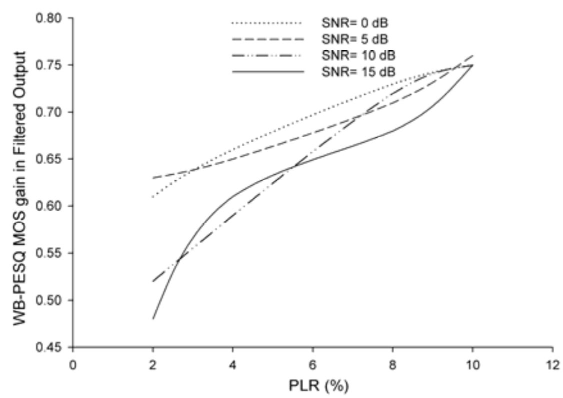
Fig.7 Signal coded with AMR-WB 18.25 kbps



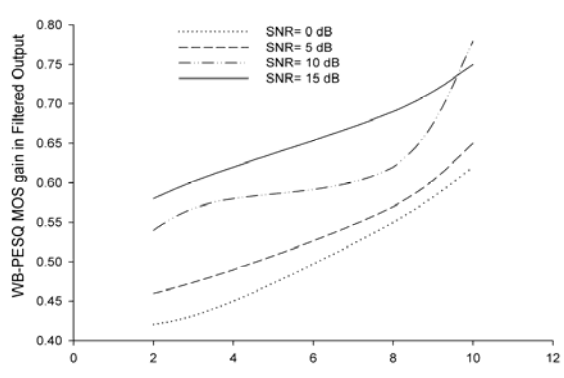
(a) Babble Noise at Packet Size = 20 ms



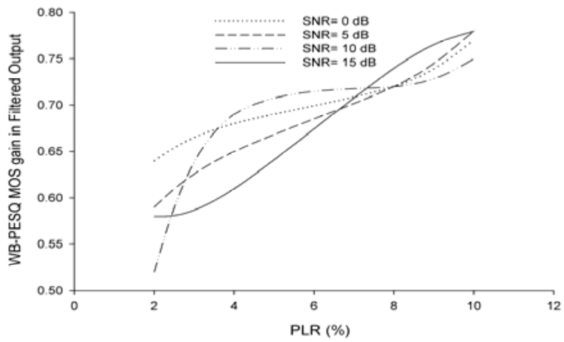
(b) Babble Noise at Packet Size = 120 ms



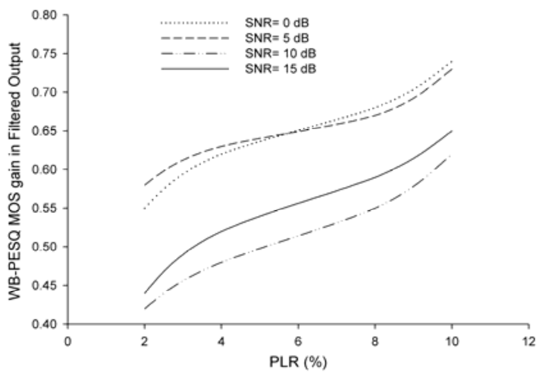
(c) Car Noise at Packet Size = 20 ms



(d) Car Noise at Packet Size = 120 ms

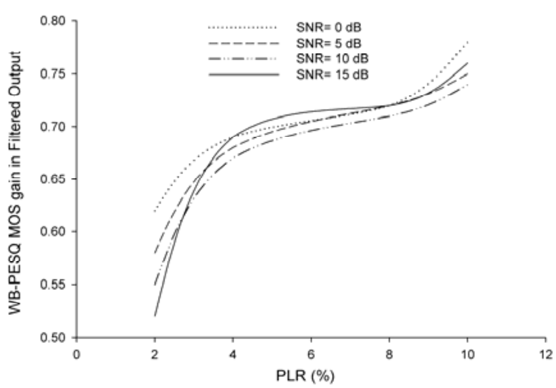


(e) Street Noise at Packet Size = 20 ms

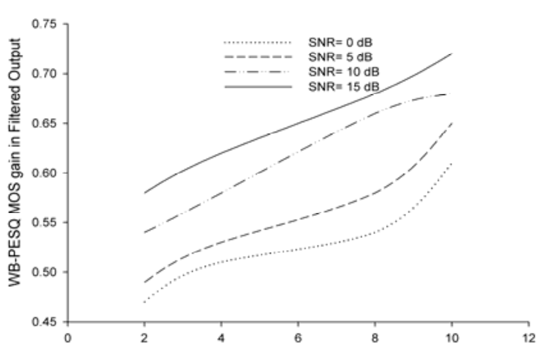


(f) Street Noise at Packet Size = 120 ms

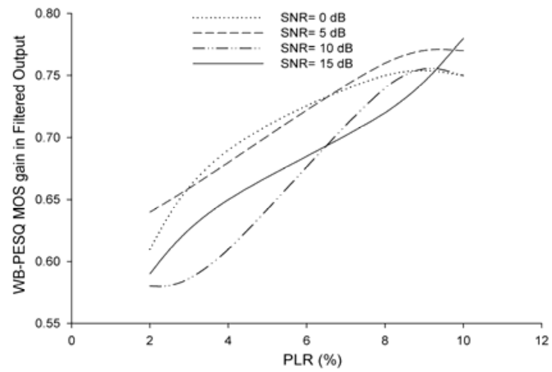
Fig. 8 Effect of IFIR filter on VoIP speech signal with AMR-WB (15.85 kbps) codec for noise at different SNR



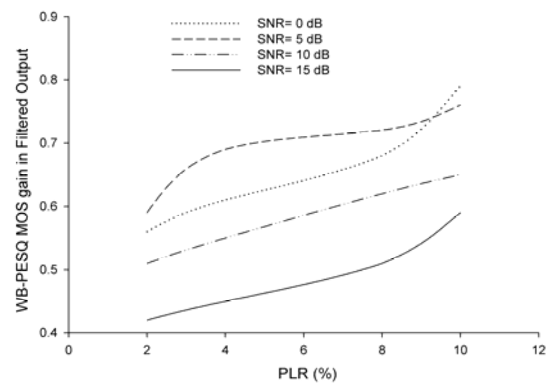
(a) Babble Noise at Packet Size = 20 ms



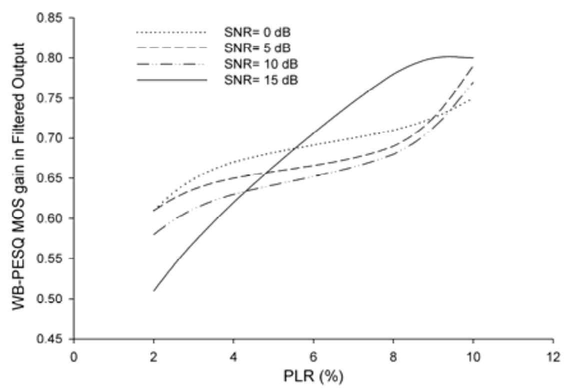
(b) Babble Noise at Packet Size = 120 ms



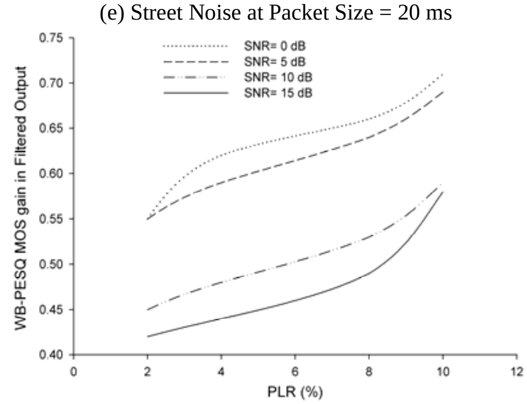
(c) Car Noise at Packet Size = 20 ms



(d) Care Noise at Packet Size = 120 ms



(e) Street Noise at Packet Size = 20 ms



(f) Street noise at Packet Size = 120 ms

Fig.9 Effect of IFIR filter on VoIP speech signal with AMR-WB (18.25 kbps) codec for noise at different

(III) Comparison with interleaving method

Interleaving technique is widely used packet loss concealment technique when the multiple frames are used in single VoIP for concealment of the lost packets during VoIP communications. The wideband coder AMR-WB is also capable for supporting the interleaving as packet loss concealment technique during the network communications [28]. To reduce the effect of packet loss on perceived speech quality, the lost packets have to be regenerated at the receiver using packet loss concealment algorithms. Goodman et.al [29] had used Waveform substitution algorithms successfully for pulse code modulation (PCM) speech coder. The lost packets were also regenerated with the use of time scale modification algorithms [30, 31]. The lost packets during network are also estimated with interleaving technique [32, 33]. The interleaving process over four consecutive frames is depicted in Fig.10.

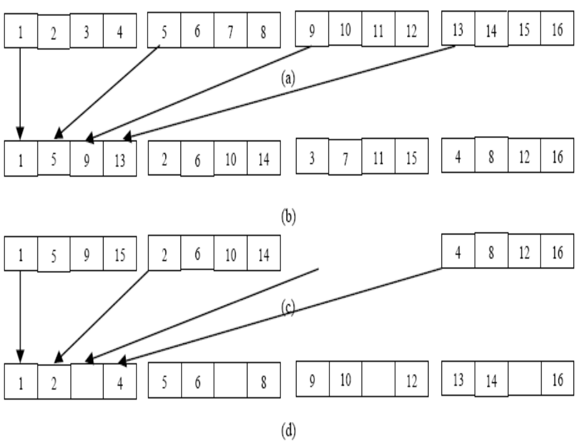
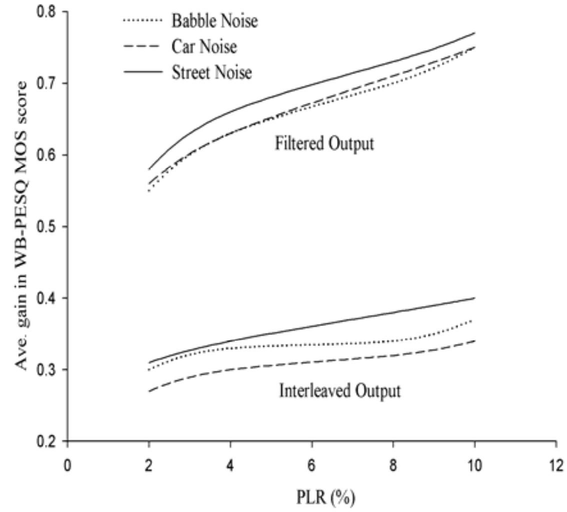


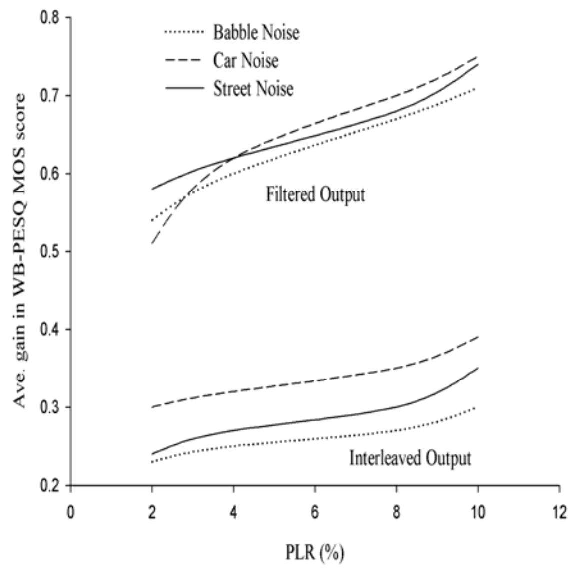
Fig. 10 Interleaving packet loss concealment scheme (a) Original four frames (b) Four frames interleaved (c) Frame loss (d) Reconstructed frames

(a) Comparison at various SNR and PLR

The average gain in WB-PESQ MOS scores is taken for various noise types at different packet loss rates for both filtered and interleaved output. The comparison results for wideband VoIP systems are presented in Table II- Table III. The significant improvement can be noticed in the results obtained with IFIR scheme in VoIP system at various network conditions. The proposed system gives much better results for small packet sizes and also outperforms the existing interleaving technique for packet loss concealment at large packet size.



(a) Packet Size = 20 ms

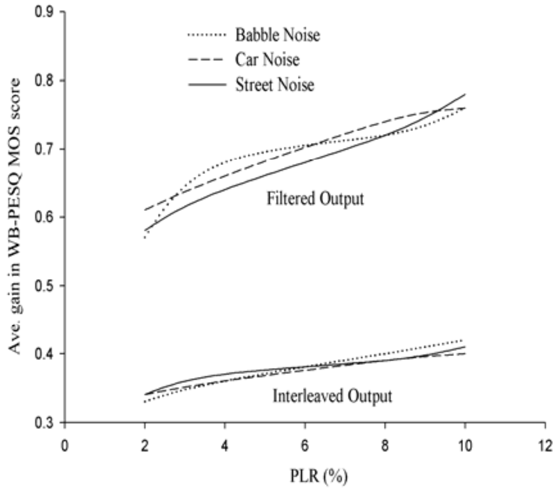


(b) Packet Size = 120 ms

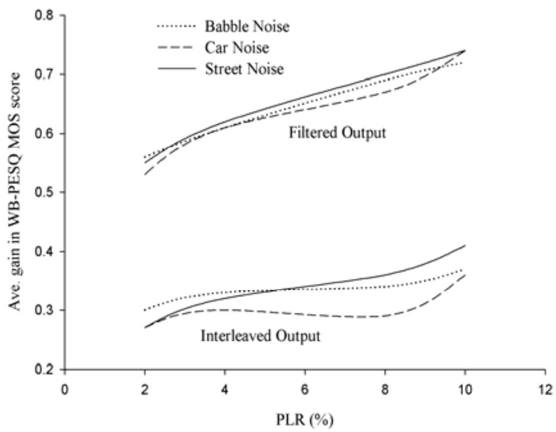
Fig.11 Comparison results for noise at different SNR for filtered and interleaved VoIP speech signal coded with AMR-WB (15.85 kbps) codec

(b) Comparison for various noise types:

The comparison results at various packet sizes with varying packet loss rates for different noise types is presented in Fig.11-Fig.12. The average gain in WB-PESQ MOS scores is taken for various SNR at different packet loss rates for both filtered and interleaved output. For wideband VoIP systems, the proposed IFIR filter results outperform the interleaved results for each type of noise used in the present work.



(a) Packet Size = 20 ms



(b) Packet Size 120 ms

Fig 12 Comparison results for noise at different SNR for filtered and interleaved VoIP speech signal coded with AMR-WB (18.25 kbps) codec

6 Conclusion

The work in this paper proposes the use of interpolated finite impulse response (IFIR) filter for noise reduction and packet loss concealment during network simulations. The proposed IFIR filter is applied in the post-processing, i.e. after the decoding of signal. The results show that the proposed IFIR filter is very much effective in all types of noise at different SNR. The performance evaluation study shows that the proposed system gives much better results in VoIP impairments including packet loss in addition to the background noise. The simulation results for VoIP system in noisy environment with AMR-WB 15.85 kbps and AMR-WB 18.25 kbps, shows the average increase of 0.67 and 0.68 resp. in PESQ-MOS scores for packet size of one voice frame in each packet. For multiple voice frames in each packet for VoIP system, the proposed system outperforms the existing packet loss concealment technique such as interleaving, when compared during VoIP simulations in noisy conditions. The average PESQ-MOS gains of 0.57 and 0.59 is achieved with AMR-WB 15.85 kbps and AMR-WB 18.25 kbps respectively by the proposed method for six voice frames in each VoIP packet in all noisy conditions. It is found that our proposed system offer much improvement as compared to the results of interleaving method where the average increase in PESQ-MOS scores were 0.33 and 0.30 for single packet and six packets in each frame respectively for AMR-WB 15.85 kbps mode. The average gain in PESQ-MOS for AMR-WB 18.25 kbps is 0.38 and 0.32 for single and six voice frames in each packet respectively. The proposed filter not only reduces the background noise but also conceals the lost packets due to network impairments to improve the speech quality of the VoIP signal. Thus our proposed IFIR scheme can be efficiently used for VoIP applications. In future, the study can be conducted in real time with digital signal processors such as TMS320C6713.

References:

[1] J W Goralski and C M Kolon, IP Telephony, 1st ed.: McGraw-Hill, 2000

[2] Jacob Benesty, M M Sondhi, and Yiteng Huang, Handbook of Speech Processing: Springer, 2008

[3] S F Boll, "Suppression of acustic noise in speech using spectral subtraction," IEEE Transactions on Acousttics, Speech and Signal processng, vol. 27, no. 2, pp. 113-120, 1979.

[4] A.Basu K.K.Paliwal, "A speech enhancement method based on Kalman filtering," in Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing, 1987, pp. 177-180.

[5] A Z R Langi, "A DSP implementation of a voice transcoder for VoIP gateways," in Proceedings of Asia Pacific conference on circuits & systems, APCC'02, vol. 1, 2002, pp. 181-186.

[6] S Han, S J HYang, and J Kim, "Noise reduction for VoIP speech codecs using modified Wiener Filter," in novations in Systems, Computing Sciences and Software Engneering, K.Elleithy, Ed.: Springer, 2007, pp. 393-397

[7] H S Malvar and D H Staelin, "Optimal FIR pre- and postfilters for decimation and interpolation of random signals," IEEE Transactions on Communication, vol. 36, no. 1, pp. 67-74, 1988.

[8] C S Wu and A Y Wu, "A novel cost effective multi-path adaptive interpolated FIR (IFIR)- based echo canceller," in Proceddings of IEEE International Symposium on Circuits and Systems, vol. 5, 2002, pp. V-453-V-456.

[9] H P Singh, S Singh, and J Singh, "Processing of VoIP Signal using TMS320C6713 in digital domain," in Proceedings of IEEE Second International Conference on

Computer Engineering and Applications, vol. 1, 2010, pp. 606-610
[10] H P Singh, S Singh, R K Sarin, and J Singh, "Analysis of FIR interpolation filter for VoIP in noisy enviroment," in Proceeding of IEEE 2nd International Conference on Computational Intelligence, Communication Systems and Networks, 2010, pp. 268-273.
[11] J Singh, H P Singh, and S Singh, "Implementation of FIR Interpolation Filter on TMS320C6713 for VoIP," in Proceedings of IEEE Second International Conference on Computational Intelligence, Communication Systems and Networks, 2010, pp. 289-294.
[12] 3GPP.TS 26.190, "Speech codec speech processing functions; Adaptive Multi-Rate -Wideband (AMR-WB) speech codec; Transcoding functions," 2009.
[13] ITU-T Recommendation. P.862.2, "Wideband Extension to Recommendation P.862 for the assessment of wideband telephone networks and speech codecs," 2005.
[14] http://www.utdallas.edu/~loizou/speech/noizeus/
[15] M Baldi and F Risso, "Efficiency of packet voice with deterministic delay," IEEE Communications Magazine, vol. 38, no. 5, pp. 170-177, 2000.
[16] M Hassan and D F Alekseevich, "Variable packet size of IP packets for VoIP transmission ," in Proceedings of the 24th IASTED International Multi-Conference on Internet and Multimedia Systems and Applications, 2006, pp. 136-141.
[17] S Jelassi, H Youssef, C Hoene, and G Pujolle, "Voicing aware parametric speech quality models over VoIP networks," in Proceedings of the Second international conference on Global Information Infrastructure Symposium, Hammamet, Tunisia, 2009, pp. 120-127.
[18] C C Wu, K T Chen, C Y Huang, and C L Lei, "An Empirical evaluation of VoIP playout buffer dimensioning in Skype, Google Talk and MSN messenger," in Proceeding of 19th International Workshop on Network and Operating System Support for Digital Audio and Video 2009, pp. 97-102.
[19] O Hohlfeld, G Rudiger, and G Halblinger, "Packet loss in real-time services: Markovian Models generating QoE impairments," in Proceedings of IEEE 16th International workshop on quality of service, (IWQoS), 2008, pp. 239-248.
[20] M S Borella, "Measurement and Interpretation of Internet Packet Loss," Journal of Communication and Networking, vol. 2, pp. 93–102, June 2000.
[21] M Yajnik, S Moon, J Kurose, and D Towsley, "Measurement and Modeling of the temporal Dependence in Packet Loss," in Proceedings of IEEE 18h Annual Joint Conference of the IEEE Computer and Communications Societies, vol. 1, 1999, pp. 345–352.
[22] W Jiang and H Schulzrinne, "QoS Measurement of Internet Real-Time Multimedia Services," Columbia University, Technical Report CUCS-015-99, Dec. 1999.
[23] S K Mitra, Digital Signal Processing: A Computer-Based Approach, with DSP Laboratory using MATLAB, 2nd ed.: Mc.Graw Hill, 2001.
[24] Y Neuvo, C Y Dong, and S K Mitra, "Interpolated finite impulse response filters," IEEE Transactions on Acoustics, Speech and Signal Processing, vol. 32, no. 3, pp. 563-570, 1984.
[25] T Saramaki, Y Neuvo, and S K Mitra, "Design of compuationally efficient interpolated FIR filters," IEEE Transactions on Circuits and Systems, vol. 35, no. 1, pp. 70-88, 1988.
[26] A Ivanov, "Case study of recent improvements in interpolated finite impulse response (IFIR) filter design methods," http://citeseerx.ist.psu.edu/viewdoc/download;jsessionid=CAF4EDC0ED5084E79DCEC1B26E3FF50C?doi=10.1.1.84.550&rep=rep1&type=pdf, EEN 436, Section C, 2007
[27] J D Proakis and D G Manolakis, Digital Signal Processing: Principles, algorithm and applications, Third edition ed.: Prentice Hall, June2000.
[28] J Sjoberg, M Westerlund, A Lakaniemi, and Q Xie, "Real-Time Transport Protocol (RTP) Payload Format and File Storage Format for the Adaptive Multi-Rate (AMR) and Adaptive Multi-Rate Wideband (AMR-WB) Audio Codecs," RFC 3267, 2002
[29] D Goodman, G Lockhart, O Wasem, and Wai Choong, "Waveform substitution techniques for recovering missing speech segments in packet voice communications," IEEE Transactions on Acoustics, Speech and Signal Processing, vol. 34, no. 6, pp. 1440 – 1448, 1986
[30] M.Roelands W.Verhelst, "An overlap-add technique based on waveform similarity (WSOLA) for high quality time-scale modification of speech," in Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP, vol. 2, 1993, pp. 554-557
[31] A.Stenger, K.B.Younes, B.Girod H.Sanneck, "A new technique for audio packet loss concealment," in Proceedings of IEEE Global Telecommunication Conference, 'Communications: The Key to Global Prosperity, 1996, pp. 48-52
[32] J L Ramsey, "Realization of optimum interleavers," IEEE Transactions on Information Theory, vol. 16, no. 3, pp. 338-345, May 1970
[33] F Merazka, "Improved packet loss recovery using interleaving for CELP-type speech coders in packet networks," IAENG International Journal of Computer Science, vol. 36, no. 3, Feb 2009.