*Sindh Univ: Res: Jour: (Sc: Ser:) Vol:31 (1) 35-38(1999)*

# PACKETIZED SPEECH QUALITY IMPROVEMENT BY POST-PITCH FILTERING APPROCH

Lubna Mustafa, A.Shah , K.R Khoumbati* and A.M.Kondoz**

Institute of Information & Technology, University of Sindh, Jamshoro, Pakistan

## Abastract

In packetized speech communication, speech packets can be corrupted due to noise or can be lost due to transmission delay. These packets cause speech quality degradation. For low bit rate speech, lost packets problem is severe, because speech coded at low bit rates is already sacrificed over bit rates reduction to save bandwidth and quality loss due to the limitation and coding algorithm limitations. Post-pitch filtering is an approach by which the effects of lost packet can be mitigated. In this paper, a Codebook Excited Linear Prediction (ECLP) system of 4.3 kb/s speech is packetized and a random packet loss phenomenon is simulated. The lost packets are reconstructed by a cascaded response of post-pitch filter and LPC filter. In this paper the simulation results for various packet loss rates and corresponding speech quality measures are presented.

**Key Words: Speechquality, post-pitch filtering,**

## Introduction

Speech coding techniques has been existing since 1960's by the introduction of Pulse Code Modulation (PCM) system of 64 kb/s in telephone. Since then speech coding and compression field has been getting a lot of attention especially for the reduction of bit rates and simultaneously for maintaining the quality, soon after PCM Adaptive Differential Pulse Code Modulation (ADPCM) was developed at 32 kb/s rate.

The quality of ADPCM coder is perceptually rated very close to PCM; whereas, the bit rate requirement is cut down by half. Although the research in this area has been almost saturated and speech coders are developed at rates as low as around 2 kb/s, the work is still being carried out through the globe on various techniques for bit rate reduction, such as Codebook Excited Linear Prediction, its variants and prototype speech coding techniques.

Experiments for packetization at higher bit rates have been facilitated through data communication networks [1]. But at low bit rate speech packetization the doors are just opened for experimentation and researching. Particularly the major obstacle for low bit rate packetization is the quality of speech. For higher bit rates, such as, 3264 kb/s, the quality does not matter much. In this the issue of lost packets, quality loss and recovery in terms of subjective measures is presented.

## Packet formation

Low bit rate coding techniques are based on a sequential coding of segments of speech. The choice of the size of these segments is parameter being used for

determining the bit rate and the quality of speech. For CELP type speech algorithm [2], three blocks were used to encode and decode each segment of size x samples the parameters of these blocks were optimized and then decoded often, by transforming them into various domains. These transformed parameters then coded and

**Cnetre for communication System Research, School of Electronics Engineering, Information Technology and Mathematics, University of Surrey, Guildford, Surrey, UK.

CamScanner

bit stream of digital codes segment by segment are transmitted. At the receiving side of the system, a similar algorithm is activated to reverse the job done at the transmitter and the particular segments are reconstructed. As part of the algorithm requirements the coding parameters of each segments are restored for coding the next incoming segments.

For any lost packet (a packet or a part of packet) these parameters of a previous frame are exploited.

**Simulation parameters**

For simulation three packet sizes are considered (a) packet of single segment of speech, in this case 30 ms.( b) packet of two segments of speech and (c) a packet of three segments of speech.

The loss of packets are simulated on a random basis which is based on the random pick of the packets. All the packets are equally probable for the loss and no priority is given to any packet.

The lost packets are reconstructed at the receiver as soon as the loss flag is activated the packet loss mechanism.
The packet reconstruction activates by employing the backward auto correlation on the previous frame parameters.

**Lost frame reconstruction**

For the loss of a single frame per packet, the previous frame parameters are exploited as follows:

The excitation vector gains for all the sub frames are set to zero which means that there is no secondary excitation or contribution from code book is zero.

The LPT delay d and gain • are calculated as follows: Firstly the cross correlation function Speech in the previous output frame before post filtering is computed as given below:

$R(t)=\sum^{80} * s ( i ) *s( i-t )$ for 19<• <80
$i = 1$

Then Delay D is set to • where R( • ) is maximum.
Compute $C( D ) = \sum^{80} s( I-D ) * s( i-D)$
$i =1$

The • is calculated by R( D ) / C( D ).
Then D and • are used for all sub frames in the current frame.

During the second consecutive frame loss, the computed D is the previous case is used for all the sub frames and the • is scaled by • * 0.8. The LPC parameters of the previous frames are retained and used. For any further loss the excitation parameters as well as the LTP • gain for all sub frames is set to zero ad remaining parameters are retained.

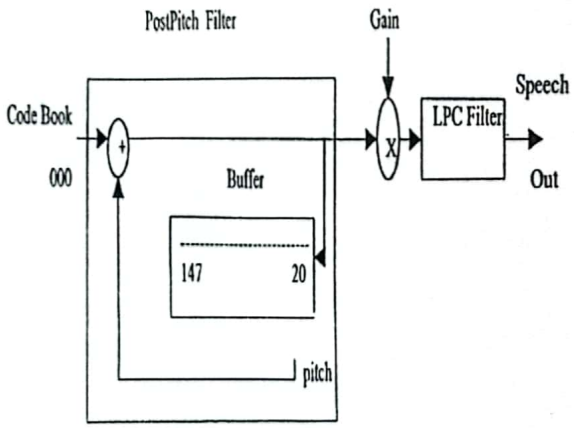



**Fig. 1 Post Pitch Filter Excitation recovery**

CamScanner

## Simulation results

The post-pitch filter is basically another form of LTP filter as shown in Fig. 1. For any lost packet the reconstruction scheme is activated and lost speech segments are filled with the reconstructed segment. The LTP filter and post-pitch filter are different in their operation, only by scaling of the pitch delay by a factor. These scaling factors cause the pitch pulse sharper and speech quality becomes smoother than the excitation is prepared form LPT filter only.

The Mean Option Score [3,4] results are presented in Fig.2., in the graph, the packet of a single frame has scored better than packet formed of double frames, and the packet of three frames has less score than both ones. Naturally, the loss of three consecutive frames can not be recovered by previous frame parameter exploitation only. The nearest and best substitute in this case is the next adjacent frame only. For the loss of three consecutive frames the distance signal has traveled from (n+0 to n+3) and is three frames away. This has caused a lot of transition and if characteristics of the signal are changed then the parameters of coding algorithm are also changed.

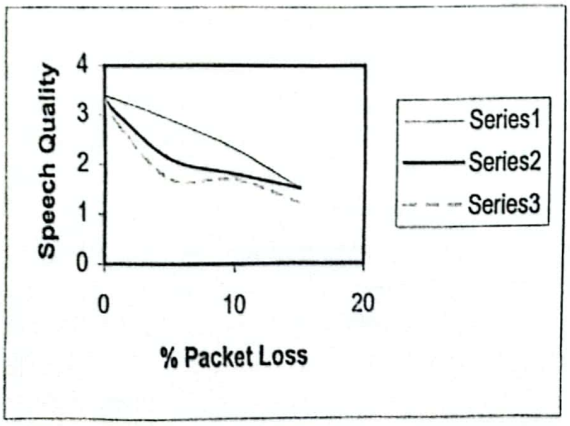



**Fig.2 MOS result for post-pitch filtering**

## Discussion

The loss of frames is considered on a random basis. The loss flag may hit any region of the speech signal. For example, the steady state where signal is periodic and transitional regions.

During transitional regions the signal energy is either building up or decaying and these areas are very sensitive and loss of the packet within this region is difficult to reconstruct.

The coding algorithm of CEL 4.3 kb/s, quality already sacarified over the bandwidth saving does not sustain additional loss of packets.

## Conclusion

The smaller size packets are better to be reconstructed due to the fact that the previous frame parameters can be a better substitute for next single frame. But the longer size packets of size of 2-3 frames the speech signal has traveled a distance of 60-90 ms apart. The reconstruction parameters of a lost frame for these packets can not be providing better substitute.

The Excitation recovery from post-pitch filter has improved the quality subjectively better than reported previously [5], where the LTP filter was the source of excitation recovery only.

## References

1. CLIFFORD J. WEINSTEN and JAMSES W. FORGIE "Experience with speech communication in packet network", IEEE Journal on selected areas in communication, Vol. S AC-No –6, Dec . 1983
2. A..M. KONDAZ "Digital speech coding for low bit rate communication systems",

CamScanner

John wiley and sons ltd . chichester westg Sussex, England, November 1995.

3. NOBUHIKO KITAWAKI, HIROMI NAGABUCHI AND KENZO ITOH..
"objective quality evaluation for low-bit-rate speech coding system ".IEE, Journal on Selected Areas of Communication Vol-6.No.2,Feb: 1988.
4. AUGUSTINEH. GRAY, JR and JOHN D.MARKEL."Distance measure for Speech 1 Processing"IEEE Transactions on Acoustics, speech and signal processing, ASSP-24,No.5, October 1976.
5. A.Shah, A.W. Ansari and Lubna Mustafa, "Low bit rate speech packetization and packet reconstruction techniques", accepted in the Mehran University Research journal of Engineering and Technology.

CamScanner