Comparison of Narrowband and Wideband VoIP using TMS320C6713 DSP Processor

Harjit Pal Singh
Department of Physics
Dr. B.R. Ambedkar National Institute of Technology
Jalandhar, India

Sarabjeet Singh
IEEE Member
Department of Physics
Dr. B.R. Ambedkar National Institute of Technology
Jalandhar, India

Jasvir Singh
Department of Electronics Technology
Guru Nanak Dev University
Amritsar, India

ABSTRACT
The speech of the Voice over Internet Protocol (VoIP) system is degraded by network layer problems which include delay, packet loss and jitter. The implementation of signal through digital signal processor can improve the quality of degraded VoIP signal. The work in this paper presents the comparison of speech quality for narrowband and wideband VoIP using TMS320C6713 DSP processor. The VoIP simulations are conducted for G.729A and AMR-WB speech coders at different packet loss rates. The digital filtering algorithm is implemented on degraded VoIP speech signal. The results of implementation experiment indicate much improvement in signal quality with wideband coders. The results are validated through the measurement of enhancement signal using perceptual evaluation of speech quality (PESQ) measurement.

General Terms
Voice over Internet Protocol, Digital Signal Processing

Keywords
Narrowband VoIP, Wideband VoIP, Signal Processing, TMS320C6713 DSP Processor

1. INTRODUCTION
Today's most demanding technology for transmission of voice, Voice over IP has the capability to facilitate voice and data convergent at an application layer [1]. But QoS is one of the problems in implementation of voice over IP. The quality of voice is the characteristic of the IP network and it is affected by parameters such as packet loss, packet delay, service capability and congestion. The network traffic and interference could cause jitter which is the variation in time between packets. This results in harmful effects on the quality of VoIP, since speech packets can be discarded when routers or gateways are congested. Due to the real time requirement for interactive speech transmission, it is usually impossible for the receivers to request the sender to retransmit the lost packets. When voice packets do not arrive before their playout time, they are considered as lost and cannot be played when they are received. One of the most difficult problems in such networks is the packet loss issue. Even a single lost packet may generate audible distortion in the decoded speech signal [2]. 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 [3, 4]. This work in this paper presents the implementation of degraded VoIP signal through digital signal processor using signal processing algorithms for making comparison of narrowband and wideband VoIP.

The next section presents the brief description of the coders used in this work. The Section III provides the description of the simulation of the VoIP network impairments. The experimental setup, filter design and implementation through TMS320C6713DSP processor is presented in Section IV. In Section V the implementation results and discussion is presented. The last section concludes the paper and present future work.

2. CODEC
2.1 G.729A coder
The G.729A coder is based on a conjugate structured: code-excited linear prediction (CS-ACELP) coding model. In this model the throat and mouth are modeled as a linear filter and voice is generated by a periodic vibration of air exciting this filter. The locally decoded signal is compared against the original signal and the coder parameters are selected such that the mean-squared weighted error between the original and reconstructed signal is minimized. The CS-ACELP coder is designed to operate with an appropriately band limited signal sampled at 8000Hz. The input and output samples are represented using 16-bit linear PCM. The coder operates on frames of l0ms corresponding to 80 samples. For each frame, the speech signal is analyzed to extract the parameters of the CELP model which include linear-prediction filter coefficients, adaptive and fixed- codebook indices and gains. At the decoder, these parameters are used to retrieve the excitation and synthesis filter parameters. The speech is reconstructed by filtering through the short term synthesis filter. After computing the reconstructed speech, it is further enhanced by post-filter [5].

2.2 AMR-WB coder
Adaptive multi-rate wideband (AMR-WB) [6, 7] was standardized by 3GPP in 2000 for use in 3G mobile systems and was also adopted by ITU in 2001 as their latest wideband coder, G.722.2. It operates at 9 varying modes of 6.6 kbps up to 23.85 kbps bit rates and describes the detailed mapping from input blocks of 320 speech samples in 16 bit uniform PCM format to encoded blocks of 132, 177, 253,285, 317, 365, 397,461 and 477 bits and from encoded blocks of 132, 177, 253,285, 317, 365, 397,461 and 477 bits to output blocks of 320 reconstructed speech samples. AMR-WB offers even lower bit rate compressions as well as the ability to quickly adapt to varying

compression as per the network conditions. The bandwidth is automatically conserved when network congestion is high. When the congestion returns to a normal mode, a lower-quality bit rate is restored.

3. IP NETWORK MODELING

The simulation of VoIP system was performed where each packet contains one frame. Packet losses are not independent on a frame-by-frame basis, but appear in bursts. The packet loss can be approximated by Markovian loss model such as Gilbert model, as discussed by Bolot in [8]. Most research in VoIP networks uses a Gilbert Model to represent packet loss characteristics [9]-[12]. Thus simulation of IP network was performed by using a 2-state Gilbert Model. The model has two states reflecting whether the previous packet is received or lost. The state "0" represents that a packet being correctly received and state "1" represents that a packet being lost. The Gilbert model is shown in Figure 1.

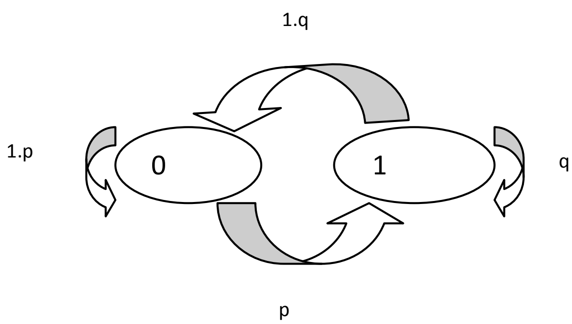


Figure.1 Two State Gilbert Model

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. The further detail about the model can be found in [13]. 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 $(p1)$.

$$p0 = \frac{q}{p+q}, \quad p1 = \frac{p}{p+q} \qquad (1)$$

The transition matrix is given as

$$P = \begin{pmatrix} 1-p & p \\ q & 1-q \end{pmatrix} \qquad (2)$$

3.1 VoIP Simulations

The VoIP simulations were carried out for making comparison between narrowband and wideband speech coders. The speech samples for simulations were taken from [14]. The speech signals are encoded with G.729A for narrowband and AMR-WB for wideband signal processing and then packetized. The network impairments were introduced into VoIP frames with the modeling of the IP network through above discussed Gilbert model. The speech signal of VoIP system is degraded at 2% to 10 % packet loss rates (PLR). The degraded output signal is then fed into the DSP processor for further processing.

4. EXPERIMENTAL SETUP

The system for the implementation of DSP algorithms on the VoIP speech signal is shown in the Figure. 2. The system comprises two main blocks, a PC application implementing the user interface, file management and control, and a DSP application running in the DSK6713, implements the speech processing algorithms. The applications exchange data through the host processor interface [15] of digital signal processor. The PC application, with its intuitive and easy-to-use GUI, enables real time monitoring, control and analysis of the real time process which is performed in the DSP processor without affecting its performance. This application manages the data and algorithms library files.

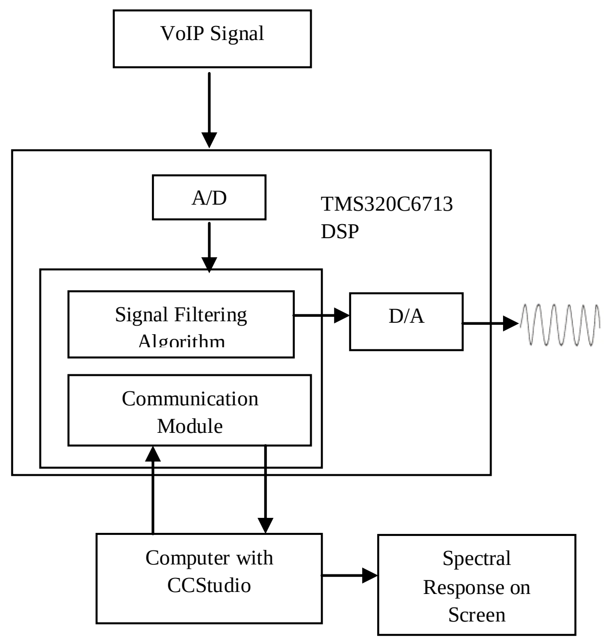


Figure.2 System Architecture

4.1 TMS320C6713 DSP Processor

The TMS320C6713 DSP composes the floating-point DSP generation in the TMS320C6000™ DSP platform. The C6713 device is based on the high-performance, advanced very-long-instruction-word (VLIW) architecture developed by Texas Instruments (TI), making this DSP an excellent choice for multi-channel and multifunction applications. The DSK also serves as a hardware reference design for the TMS320C6713 DSP. Operating at 225 MHz, the C6713 delivers up to 1350 million floating-point operations per second (MFLOPS), 1800 million instructions per second (MIPS), and with dual fixed-/floating-point multipliers up to 450 million multiply-accumulate operations per second. An on-board AIC23 codec allows the

DSP to transmit and receive analog signals. McBSP0 is used for the codec control interface and McBSP1 is used for data. Analog audio I/O is done through audio jacks that correspond to microphone input, line input, line output and headphone output. A programmable logic device called a CPLD is used to implement glue logic that ties the board components together. The CPLD has a register based user interface that lets the user configure the board by reading and writing to the CPLD registers [16]. Code composer studio (CCStduio) IDE is an integrated development tool that is a key component of the "eXpressDSP Software and Development Tools" strategy of Texas Instrument. CCStudio includes the TMS320C67xx code generation tools along with the APIs and plug-ins for both DSP/BIOS and RTDX. Code Composer communicates with the DSK through an embedded JTAG emulator with a USB host interface [17]. The schematic diagram of TMS320C6713 DSP processor is given in Figure.3.

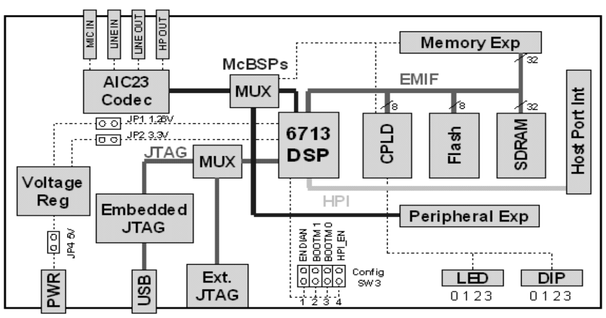


Figure.3 TMS320C6713 DSP Processor

4.2 Filter design and Implementation

To improve the performance of the VoIP speech signal, the window based FIR filter is designed. FIR filter is chosen because it is particularly useful for applications where exact linear phase response is required. FIR filters are filters having a transfer function of a polynomial in z- and is an all-zero filter in the sense that the zeroes in the z-plane determine the frequency response magnitude characteristic. The z transform of an N-point FIR filter is given by

$$H(Z)=\sum_{n=0}^{N-1}h(n)Z^{-n} \qquad (3)$$

The major advantages of using window method are their relative simplicity as compared to other methods and ease of use. The fact that well defined equations are often available for calculating the window coefficients has made this method successful. The Kaiser window is used to design the FIR filter. The Kaiser window with parameter β is given as

$$W(n)=\begin{cases}\dfrac{I_o(\beta)\sqrt{1-(2(n+1)(N+1))^2}}{I_o(\beta)} & n=0,1\ldots N\\ 0, otherwise\end{cases} \qquad (4)$$

The Bartlett window reduces the overshoot in the designed filter but spreads the transition region considerably. The Hanning, Hamming and Blackman windows use progressively more complicated cosine functions to provide a smooth truncation of the ideal impulse response and a frequency response that looks better. The best window results probably come from using the Kaiser window, which has β, which allows adjustment of the compromise between the overshoot reduction and transition region width spreading [18]. The proposed FIR filter is designed in MATLAB and then this filter is implemented to degraded VoIP signal through DSP processor. The window based FIR filter is programmed in C-language for its implementation on DSP TMS320C6713 processor [19]. The Figure.4 shows the programming interface of the code composer studio.

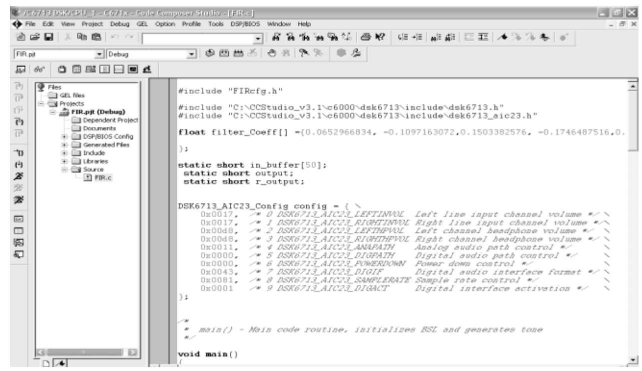


Figure.4 Programming Interface of Code Composer Studio

5. IMPLEMENTATION RESULTS

This section presents the results of narrowband and wideband VoIP signal processing through digital signal processor. The degraded signal is passed through ADC of then DSP processor and then filter operation was done on the signal and the resultant enhanced speech signal is taken from DAC of DSP board. The various results of the DSP processor are presented here:

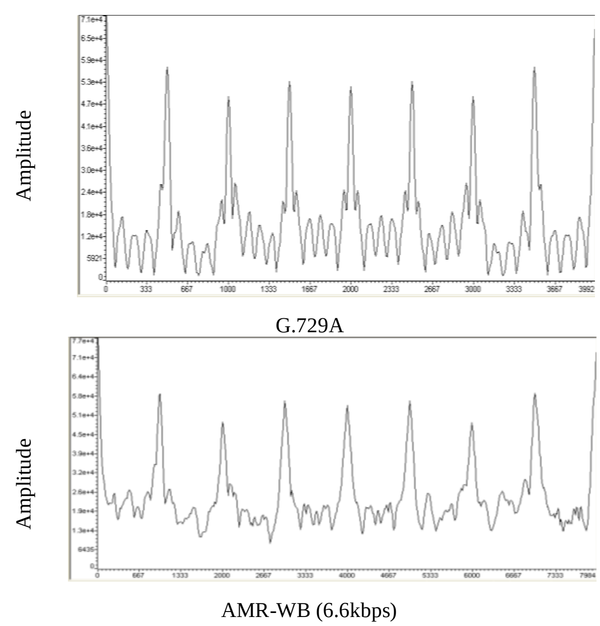
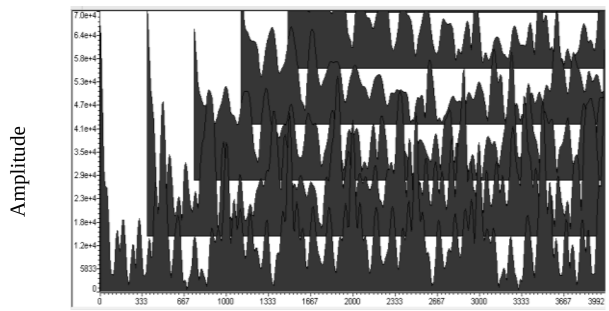
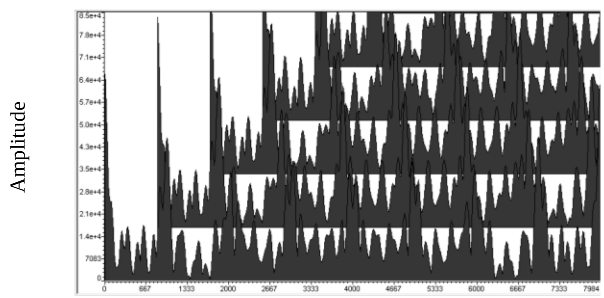


Figure.5 FFT Magnitude of VoIP Signal

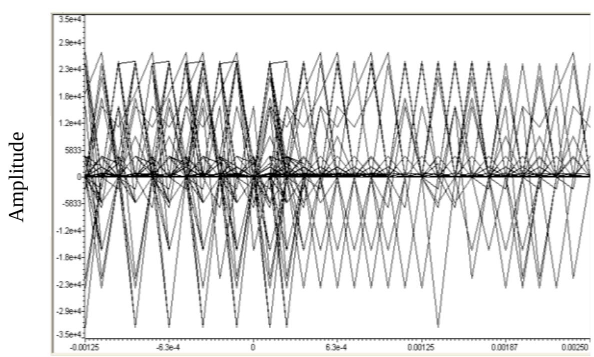


G.729A

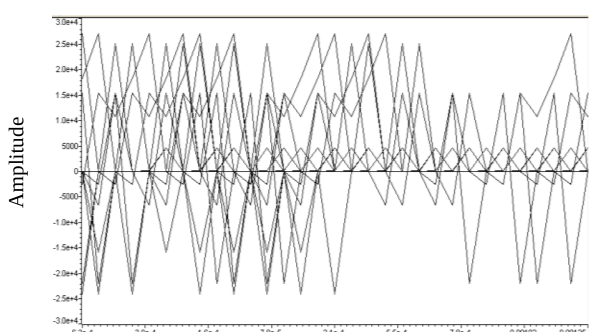


AMR-WB (6.6kbps)

Figure.6 Waterfall FFT Magnitude of VoIP Signal



G.729A



AMR-WB (6.6kbps)

Figure.7 Eye Diagram Analysis of VoIP Signal

5.1 Discussion

The spectral analysis of the enhanced VoIP signal was done through FFT magnitude, waterfall analysis and Eye diagram analysis as given in Fig 5-Fig.7. The spectral analysis of degraded VoIP signal indicates that the signal quality can be can be improved with efficient digital signal processing algorithms with wideband speech coders in VoIP applications. The smoothness in the FFT and waterfall analysis results in Fig. 5 and Fig.6 indicates improvement in signal quality for wideband VoIP signal in comparison to the narrowband VoIP signal. The residual noise is clearly visible in eye diagram results for narrowband VoIP signal and this noise does not have much effect on wideband VoIP signal, as depicted in Fig.7. The filter was designed in such a way that it introduces low delay and produces the improved speech signal. The filter also reduce the background noise which was added while transmission through network. The experiment was performed on different packet loss rates but the spectral analysis results of VoIP signal at packet loss rate of 2% are shown in this paper. The results are also validated with other packet loss rates. The performance is evaluated with Perceptual Evaluation of Speech Quality (PESQ) measurement defined by ITU-T recommendation P.862 [20] for narrowband coders and wideband extension- Perceptual Evaluation of Speech Quality (WB-PESQ) measurement defined by ITU-T recommendation P.862.2 [21] for wideband coders. After comparing the degraded signal with the original one, the PESQ measurement gives the subjective measurement as Mean Opinion Scores (MOS) value, as presented in Fig.8 and Fig.9 for G.729A and AMR-WB coder. The mean opinion scores (MOS) for the filtered output wideband coder is approximately near to 4.0 at different packet loss rates, which is very well accepted for signal quality.

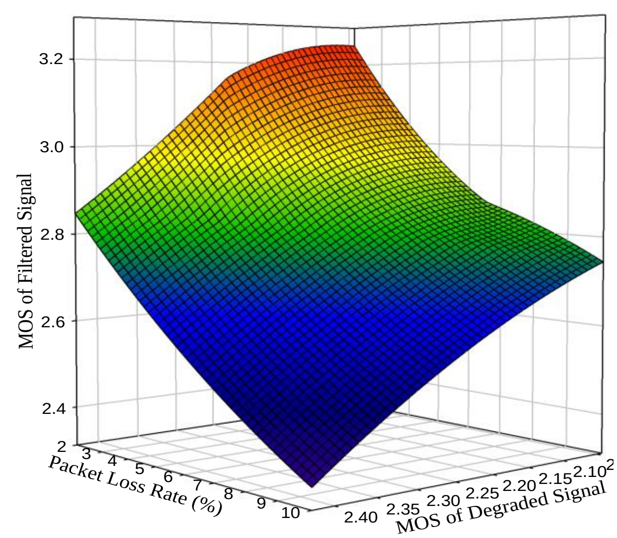


Figure.8 Packet loss rate vs. MOS for G.729A

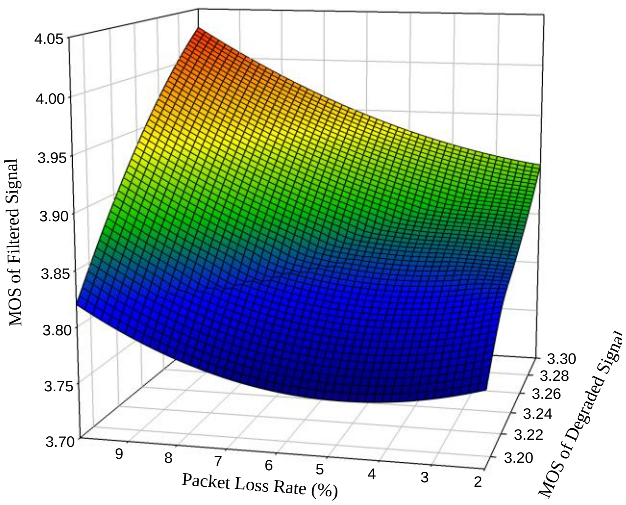


Figure.9 Packet loss rate vs. MOS for AMR-WB

6. CONCLUSION AND FUTURE WORK

The signal processing of degraded narrowband and wideband VoIP speech was performed in this work. The real time implementation through DSP processor of VoIP speech signal not only improves the speech quality but also try to retain the spectral shape of original signal. The comparison results presented in this paper indicate that the spectral quality of VoIP speech signal is much improved with wideband coders. In future the study can be used for improving the speech quality using various signal processing algorithms performed at higher frequency digital signal processors.

7. ACKNOWLEDGMENTS

The authors are thankful to Ministry of Human Resource and Development of India for carrying out this work.

8. REFERENCES

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

[2] Singh, H.P., Singh, S., and Singh, J.,2010, Computer modeling and performance analysis of VoIP under different strategic conditions, In Proceedings of IEEE Second International Conference on Computer Engineering and Applications, pp.611-615.

[3] Varshney, U., Snow, A., McGivern, M., and.Howard, C.,” Voice over IP”, Communications of the ACM, Vol.45, No. 1, 2002, pp. 89-96

[4] Singh, H.P., Singh, S., and Singh, J., 2010, Processing of VoIP signal using TMS320C6713 in digital domain, In Proceedings of IEEE Second International Conference on Computer Engineering and Applications, pp.606-610.

[5] ITU-T Recommendation, 1996, G.729 Annex A. Reduced complexity 8 kbit/s CS-ACELP speech codec

[6] 3GPP TS 26.171,2008, Speech codec speech processing functions; Adaptive Multi-Rate - Wideband (AMR-WB) speech codec; General description

[7] Bessette, B., Salami, R., Lefebvre, R., Jelinek, M., and co-authors, “The adaptive multirate wideband speech codec (AMR-WB)”, IEEE transactions on speech and audio processing, vol. 10, No. 8, Nov 2002, pp.620-636

[8] Bolot, J.C.,1993, End-to-end frame delay and loss behavior in the internet, In Proceedings of ACM SIGCOMM, pp.289-298

[9] Sannech, H., and Le, N.T.L.,2000, Speech property-based FEC for Internet Telephony applications, In Proceedings of the SPIE/ACM/SIGMM Multiledia Computing and Networking Conference, pp.38-51

[10] Jelassi, S., Youssef, H., Hoene, C., and Pujolle, G., 2009, Voicing aware parametric speech quality models over VoIP networks**,** In Proceedings of the Second international conference on Global Information Infrastructure Symposium, Hammamet, pp. 120-127

[11] Wu, C.C., Chen, K.T.,Huang, C.Y., and Lei, C.L.,2009, An Empirical evaluation of VoIP playout buffer dimensioning in Skype, Google Talk and MSN messanger, In Proceedings of NOSSDAV’09, pp-97-102

[12] Singh, S., Singh, H. P., and Singh, J.,2010, Spectral Analysis of Speech Quality in VoIP for G.729A andAMR-WB Speech Coders, In Proceedings of IEEE 2nd International Conference on Computational Intelligence, Communication Systems and Networks,pp. 182-187

[13] Hohlfeld, O., Rudiger.G, and Halblinger, G.,2008, Packet loss in real-time services:Markovian Models generating QoE impairments, In Proceedings of 16th International workshop on quality of service, pp:239-248

[14] Open Speech Repository. http:// www.voiptroubleshooter.com/open_speech/.

[15] Texas Instrument 2003, TMS320C6000 DSP Host Port Interface (HPI) Reference Guide

[16] Spectrum Digital Incorporated, 2004, TMS320C6713 DSK Technical Reference

[17] Texas Instruments, 2005, Code Composer Stodio User’ Guide, Texas Instruments

[18] Proakis, J.D., and Manolakis, D.G., 2000, Digital Signal Processing :Principles, algorithm and applications, Third edition.

[19] Texas Instruments, August 2002, TMS320C6000 Programmer’s Guide

[20] ITU-T Recommendation, 2001, P.862: Perceptual evaluation of speech quality (PESQ); an objective method for end-to-end speech quality assessment of narrowband telephone networks and speech codec

[21] ITU-T Recommendation, 2005, P.862.2: Wideband extension to recommendation P.862 for the assessment of wideband telephone networks and speech codecs