ISSN (Print) : 2319-5940
ISSN (Online) : 2278-1021

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 2, Issue 4, April 2013

Near Perfect Reconstruction of Speech Signal Using Filter Bank Techniques

Prashant Upadhyaya[1], Reetu Singh Kaurav[2], Rashi Sharma[3], Minakshi Dubey[4], Pragya Gupta[5]

Assistant Professor, EC Department, SRSITS, Datia, Madhya Pradesh, India [1]
B.E (IVth), EC Department, SRSITS, Datia, Madhya Pradesh, India [2]
B.E (IVth), EC Department, SRSITS, Datia, Madhya Pradesh, India [3]
B.E (IVth), EC Department, SRSITS, Datia, Madhya Pradesh, India [4]
B.E (IVth), EC Department, SRSITS, Datia, Madhya Pradesh, India [5]
upadhyaya.prashant@rediffmail.com[1]

Abstract: In this paper a near perfect reconstruction of speech signal algorithms are used for reconstructing the original signal based on subband filters. In today's scenario application of voice signal has been predominate in the field of technology for security purpose or for controlling the device. This paper deals with the optimum design of filter bank for subband coding system. Also the study of perceptual speech quality is done so that the reconstructed signal should not be degraded below the desired MOS-scale.

Keywords: Subband, MOS, SBC, Analysis and Synthesis bank, SNR

I. INTRODUCTION

Today the world is been surrounded by the handheld device which are integrating with voice-related applications (e.g., voice based search directory in mobile phone and security system in laptops) based on the speech signal. On the other hand the power consumption, growing need for bandwidth conservation and enhanced privacy in these devices has been become popular [1].

From several applications almost more than 95% of applications require the speech signal to be in digital format due to flexibility it provide during processing of data, storage and opportunities for encryption. The main objective of speech coding is to represent speech with a minimum number of bits while maintaining its perceptual quality.

Hence Speech Coding or Speech Compression can be consider as the technology for obtaining compact digital representations of voice signals for the purpose of efficient transmission or storage[2].

On the other hand subband coding (SBC) [3] is a frequency domain coding technique in which the input signal is decomposed into a number of subbands so that each of these frequency bands can be encoded separately. The main objective is to reduce the effect of quantizing noise due to coding and therefore to improve the quality of speech coding systems [4].

Encoding in subband offers several advantages that can be effectively used to achieve noise reduction [1, 3]. The main advantages of this approach are the following:

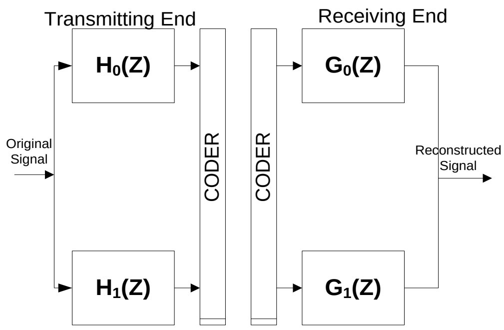


Figure 1: Analysis and synthesis bank.

(a) the quantizing noise that is generated in a particular subband is limited to that band in the reconstruction, without being allowed to spread to other bands where there may be less signal energy.

(b) bit resources can be allocated in such a way so that the number of quantizer levels and hence the reconstruction error variance can be separately controlled in each band.

The important thing that is kept in mind during the analysis on the signal is to preserve the low-frequency content with more accuracy than high-frequency bands, given limited bit resources [5].

More often the coding method used are especially designed for the speech signal which been converted into digital format. As known that speech has the range from 300 Hz to 3300 kHz

ISSN (Print) : 2319-5940
ISSN (Online) : 2278-1021

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 2, Issue 4, April 2013

hence it is generally a y band limited to 4 kHz (or 3.2 kHz) and sampled at 8 kHz.
For telecommunication purpose the simplest non-parametric coding technique is Pulse-Code Modulation (PCM) which is simply a quantizer of sampled amplitudes [1]. On the other hand speech coded at 64 kbps using logarithmic PCM is

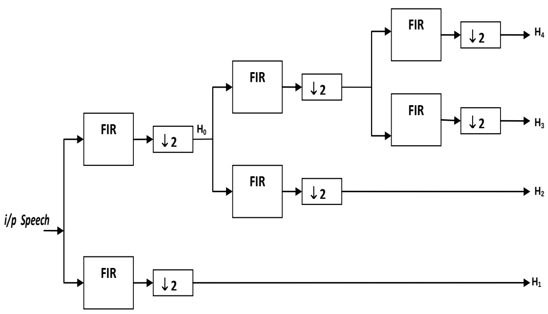


Figure 2: Tree structure filter bank of analysis bank

considered as "non-compressed" and is used as a reference for comparisons.

In this paper, we have design the subband coder for speech signal which exhibit the QMF (Quadrature Mirror Filter) properties [6]. The input signal $[x(t)]$ is down sample at the transmitting end and passed through the filter's banks and at the receiving end the signal $[\tilde{x}(t)]$ is reconstructed by up sampling the signal.

II. Analysis and Synthesis Bank

Generally the coding of speech is done at medium-rates and below and this is achieved using an analysis-synthesis process as shown in Figure1. In the analysis stage, speech is represented by a compact set of parameters which are encoded efficiently [6]. In the synthesis stage, these parameters are decoded and used in conjunction with a reconstruction mechanism to form speech.
It can be easily observed from figure 1 that the original signal band is been split into frequency sub-band using bank of filter i.e. LPF (low pass filter) H_0 (Z) and HPF (high pass filter) H_1 (Z) filter bank and then passed through the coder at the

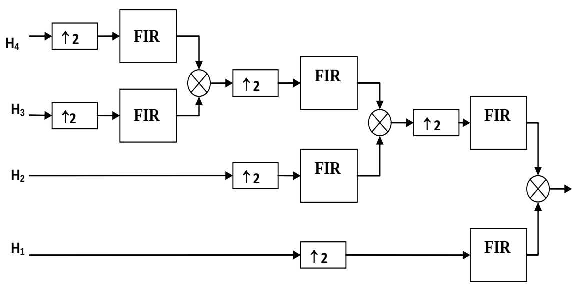
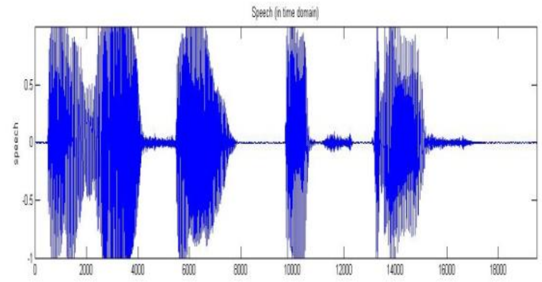


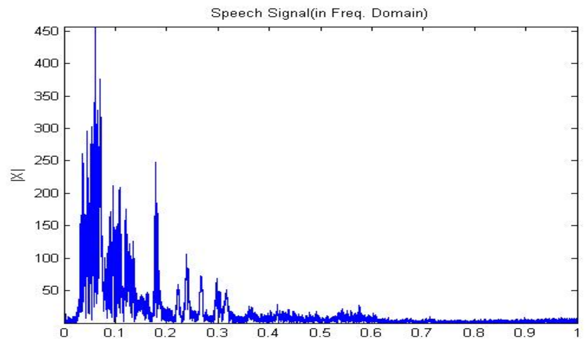
Figure 3: Synthesis bank for reconstructing signal.

transmitting end and the same process is repeated at the receiving end. The reconstructed signal is obtained which resemble the original signal [4, 6].

Analysis can be open-loop or closed-loop. In closed-loop analysis, the parameters are extracted and encoded by minimizing explicitly a measure (usually the mean square) of the difference between the original and the reconstructed



(a)



(b)

Figure 4 (a) Speech signal in time domain and (b) Speech signal in frequency domain.

speech [1].

Therefore, closed loop analysis incorporates synthesis and hence this process is also called analysis by synthesis [2]. The main motive of the sub-band coder is to produce perceptually intelligible speech without necessarily matching the waveform hence they are capable of operating at very-low rates but also tend to produce speech of synthetic quality [3, 6].

.

III. Experimental Approach

Sub band processing is based on splitting the frequency range into M segments (subbands), which together encompass the entire range. Each subband is processed independently, as called for by the specific application. The subbands are recombined after processing, to form an output signal whose bandwidth occupies the entire frequency range.

ISSN (Print) : 2319-5940
ISSN (Online) : 2278-1021

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 2, Issue 4, April 2013

Here we have used tree structure filter bank for decomposing the speech signal i.e. splitting of a signal into 2^L channels, is

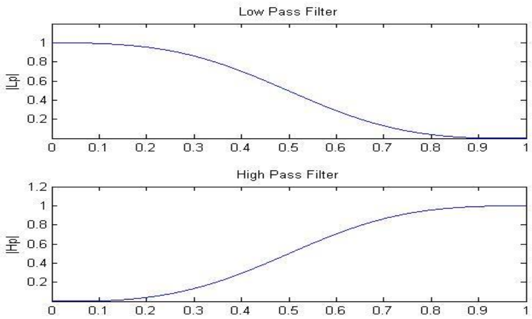


Figure 5: Filter response of the LPF and HPF

possible for any integer L using a tree structured filter bank.

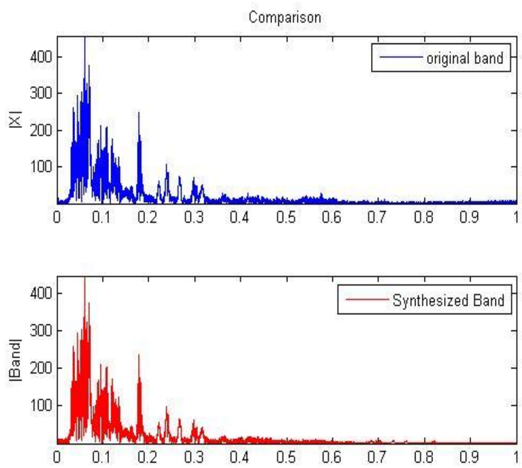


Figure 6 shows the comparisons of original signal and synthesized signal.

The upper output of the two outputs of the first level is passed to a two channel bank at a second level. This process can be continued according to the desired number of subbands as shown in figure 2. On the other hand at the receiving end i.e. synthesis banks all the outputs from the analysis bank are combined to reconstruct the original signal as shown in figure 3.

The synthesis system performs the inverse of the analysis operation. Between every two samples in each band we put in the value zero to increase the sample rate. Then we merge the two lowest bands, for example: 0-1000 Hz and 1000-2000 Hz, into the reconstructed 0-2000 Hz band.

The same operation is done to the other two bands, 2000-3000 Hz and 3000-4000 Hz, are merged into one band of 2000-4000 Hz. Again, we increase the sample rate in these two bands and merge them again. Figure 4 (a) shows the original speech signal in time domain and (b) shows the original speech signal in frequency domain. The signal is passed through the LPF and HPF filter having the filter response shown in figure 5. The process is followed for completed analysis bank. At the receiver end the inverse operation is performed i.e. the synthesis of the signal is done as shown in figure 3 and all the output of the filter are combine together to form a single signal i.e. combination of the signal obtained after passing through the filter bank. Finally we have obtained the response of the original speech signal and reconstructed speech signal, which resemble the original signal. Figure 6 shows the comparison between the original speech signal and synthesized speech signal.

IV. CONCLUSION

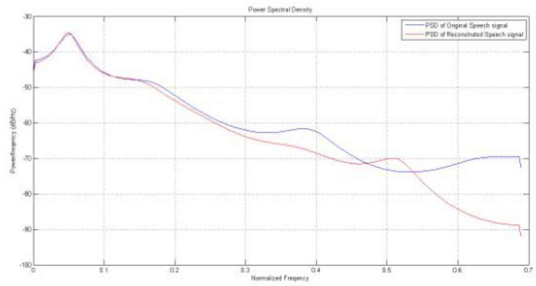


Figure 7 Power spectral densities for the original signal and reconstructed signal.

From the above results we conclude that our design procedure reconstructed the signal almost same as that of original signal. Also we have calculate the power and energy of the signal and there was only the little difference in the original signal and reconstructed signal. Figure 7 shows the power spectral density of the original signal and the reconstructed signal. Hence it can be observed from psd of these two signal that there is the change in high frequency region where as the lower frequency region is preserved, which should be exactly the case which has to be kept in mind while designing the filter. Since the most of the information is present in the low frequency region hence we have design the optimum filter for speech signal.

ACKNOWLEDGMENT

The author would like to thanks to Shri Rawatpura Sarkar Institute of Technology & Science-Datia, Madhya Pradesh for giving us the opportunities and giving us the research oriented environment in the college campus. We would also like to thanks our Associate Director Shri Sanjeev Gupta and our Principal Dr. Dheeraj Shivahare and HOD-EC department Associate Prof. Shashikant Verma for their continuous support and guidance they provide us during the entire project work.

ISSN (Print) : 2319-5940
ISSN (Online) : 2278-1021

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 2, Issue 4, April 2013

REFERENCES

[1] Speech Coding: A Tutorial Review, Andreas S. Spanias, Proceedings of the IEEE. Vol. 82, No. 10. October 1994

[2] L. Rabiner and B. Gold, Theory and Application of Digital Signal Processing. Prentice-Hall Inc., Englewood Cliffs, N. J., 1974.

[3] R. E. Crochiere, S. A. Webber, and J. L. Flanagan, "Digital coding of speech in sub-bands," Bell System Technical Journal, vol. 55, pp. 1069-1085, October 1976.

[4] Improving The Filter Bank Of A Classic Speech Feature Extraction Algorithm Mark D. Skowronski and John G. Harris IEEE Intl Symposium on Circuits and Systems, Bangkok, Thailand, vol IV, pp 281-284, May 25 - 28, 2003, ISBN: 0-7803-7761-31.

[5] Mel Filter-Like Admissible Wavelet Packet Structure for Speech Recognition O. Farooq and S. Datta IEEE SIGNAL PROCESSING LETTERS, VOL. 8, NO. 7, JULY 2001.

[6] J. D. Johnston, "A filter family designed for use in quadrature mirror filter banks," in Proceedings IEEE International Conference on Acoustics, Speech, and Signal Processing, April 1980, pp. 291-294.

Biography

Prashant Upadhyaya received the B.Tech (Electronics and Communication) degree from Uttar Pradesh Technical University, Lucknow, India in 2006 and M.Tech (Communication and Information System) degree from Aligarh Muslim University, Aligarh, India in 2012. He is presently working as Assistant Professor in SRITS-Datia. His area of research is Audio Visual Speech Recognition (AVASR), Image Processing and Digital Signal Processing. He has published many research papers in field of AVASR and other.

Reetu Singh Kaurav is the final year student of B.E (Electronics and Communication) RGPV-Bhopal and has carried out this work for the completion of their final year project. Her area of interest is Digital signal processing.

Minakshi Dubey is the final year student of B.E (Electronics and Communication) RGPV-Bhopal and has carried out this work for the completion of their final year project. She has recently got the achievement by scoring good percentile in GATE-2013. Her area of interest is VLSI and Signal Processing.

Rashi Sharma is the final year student of B.E (Electronics and Communication) RGPV-Bhopal and has carried out this work for the completion of their final year project. Her area of interest is Digital signal processing.

Rashi Sharma is the final year student of B.E (Electronics and Communication) RGPV-Bhopal and has carried out this work for the completion of their final year project. Her area of interest is Multimedia processing.