International Journal of Advanced Research in Electronics and Communication Engineering (IJARECE)
Volume 4, Issue 4, April 2015

SUB-BAND CODING OF NOISY SPEECH SIGNALS USING DIGITAL SIGNAL PROCESSING

Lalitha R Naik[1], Devaraja Naik R L[2]

ABSTRACT:

A variety of techniques have been developed to efficiently represent speech signals in digital form for either transmission or storage. The most of the speech energy is contained in the lower frequencies; we would like to encode the lower-frequency band in more bits than the high-frequency band. This paper presents a very low bit rate speech coder based on sub-band coding (SBC) is a method where the speech signal is subdivided into several frequency bands and each band is digitally encoded separately. The Audible frequency spectrum 20Hz – 20 KHz is divided into frequency sub-bands using a bank of finite impulse response (FIR) filter. The output of each filter is then sampled and encoded. At the receiver, the signals are de-multiplexed, decoded and demodulated and then summed to reconstruct the signal. Cut-off frequency at rate 4 KHz and order of the filter is 20. Noise is the major constraint with the speech signal and will consider some different type of noise, for example it could be AWGN, impulse, street noise, fading babble and more. We will implement a system to get back the de-noisy signal and for the different types of noises to be represented and implemented on GUI to ease the user to operates.

I. INTRODUCTION

In telecommunication systems that transmit and receive different types of signals e.g., teletype, facsimile, speech, video, etc., there is a requirement to process the various signals at different rates [1]. The process of converting a signal from a given rate to a different rate is called sampling rate conversion. In turn, systems that employ multiple sampling rates in digital signal processing are called multi rate digital signal processing (DSP) systems [3]. Sampling rate conversion of a digital signal can be accomplished in one of two methods. The SBC depends on a phenomenon of the human hearing system called masking. Normal human ears are sensitive to a wide range of frequencies. However, when a lot of signal energy is present at one frequency, the ear cannot hear lower energy at nearby frequencies. We say that the louder frequency masks the softer frequencies. The louder frequency is called the masker [4]. Strictly speaking, what we're describing here is really called simultaneous masking (masking across frequency).

Lalitha R Naik[1], Assistant Professor Department of computer science, karnatsak university, karnatak science college Dharwad.
Devaraja Naik R L[2], Lecturer, Dept. Of Computer Science, Karnatak science college, Dharwad

There are also non-simultaneous masking (masking across time) phenomena, as well as many other phenomena of human hearing [1].

The basic idea of SBC is to save signal band width by throwing away information about frequencies which are masked. The result won't be the same as the original signal, but if the computation is done right, human ears can't hear the difference [2].

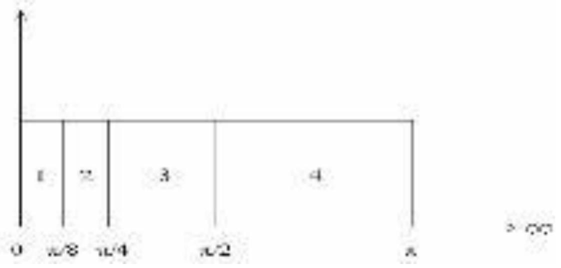


Fig 1. Frequency division principle

An example of a frequency subdivision is shown in the Figure 1. Let us assume that the speech signal is sampled at a rate Fs samples per second. The first frequency subdivision splits the signal spectrum into two equal width segments, a low pass signal (0< F <Fs/4) and a high pass signal (Fs/4 < F <Fs/2). The second frequency sub division splits the low pass signal from the first stage into two equal bands, a low pass signal (0 < F<Fs/8) and a high pass signal (Fs/8 < F <Fs/4). Finally, the third frequency subdivision splits the high pass signal from the second stage into two equal bandwidth signals. Thus the signal is subdivided into 4 frequency bands, covering 3 octaves, as shown in the figure .Decimation by a factor of 2 is performed after frequency subdivision. By allocating different number bits per sample to the signals in the 4 sub-band, we can achieve a reduction in the bit rate of the digitalized speech signal [2].

Noise is the major breach with respect to speech signal analysis. In this paper study about different types of noise is carried out. Here in this particular paper, type of the noise used for the case study is “High amplitude noise (impulse like noise)” and reduction technique is “Thresh holding technique”.

II. BASIC APPROACHES TO DIGITAL FILTER DESIGN

In case of an IIR filter design, the most common practice is to convert the digital filter specifications to analog LP prototype filter specifications, to determine the analog LPF transfer function Ha(S) meeting these specifications and then to transform it into the desired digital

ISSN: 2278 – 909X

filter transfer function H(Z). This approach has been widely used for many reasons.

a) Analog approximation techniques are highly advanced.

b) They usually yield closed form solutions.

c) Extensive tables are available for analog filter design.

Unlike the IIR digital filter design, the FIR filter design does not have any connection with the design of analog filters. The design of FIR filters is therefore based on direct approximation of specified magnitude response, with the often added requirements

H (Z) =∑ =0−1 h[ ] − …………………… (1)

To ensure a linear phase design the condition must be satisfied.

h(− 1 −) = h[ ] − h[ ]……….. (2)

Two direct approaches to the design of FIR filters are the truncated Fourier series approach and the frequency sampling approach.

a) Finite impulse response (FIR) filter

These filters are a simple class of filters that performs much of the work of digital signal processing. They are composed of multipliers, adders, and delay elements. Multipliers are rather complex parts, so sometimes they may be replaced by simpler parts. For example, to multiply by 8,it is easier to shift the number 3 places to the left(since 2^3=8). The delay elements can be thought of as registers, in fact, since registers are used to store a value for a short period of time, a register actually implements a delay element.

b) Structure of FIR filter

Describing FIR filters by equations, and how the unit impulse function works with them, for K +1 filter coefficients. (There are K + 1 of them because we start at 0 and count to K.) The number of filter coefficients is also called the number of taps. By convention, the number of taps equals the number of filter coefficients. So a filter with coefficients (b0, b1...bK) has K + 1 taps, since there are K + 1 total filter coefficients. However, it is said to be of order K. In other words, the order of the filter and the taps express the same idea, but with a difference of 1. It is possible to determine the output. It is also possible to determine an equation for the output, which is referred in the equations 4 and 5.

y[n] = b[0]x[n - 0] + b[1]x[n - 1] + b[2]x[n - 2] + ::: + b[K]x[n - K];…………………...(3)

[ ] = ∑ =0 [ ] [−]………… (4)

c) Low pass filter

The low pass allows only low frequency components to pass through it. The output of low pass filter is free of high frequency components. For the frequency response refer fig 2.

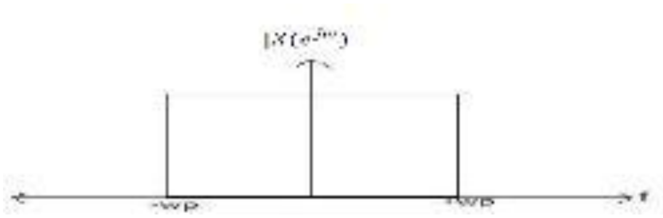


Fig 2. Required response of LPF

Generalized equation is given by the equation [4].

X (n) =sinc(wp*n)……………………… [4]

This is the required function of a low pass filter [3].

d) High pass filter

High pass filter allows only high frequency components to pass through it. It removes the all the low frequency component from signal. For the response of high pass filter, refer Fig 3.

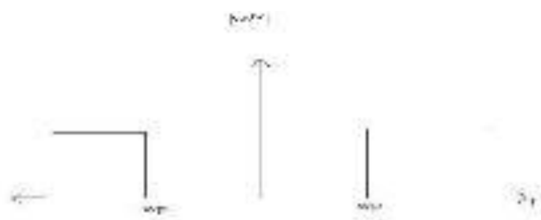


Fig 3. Response of high pass filter

Generalized equation is given equation [5]

( ) = sin(*) * − .... (5)

This is the required function to compute the high pass filter [3].

III.IMPLEMENTATION

TRANSMITTER:

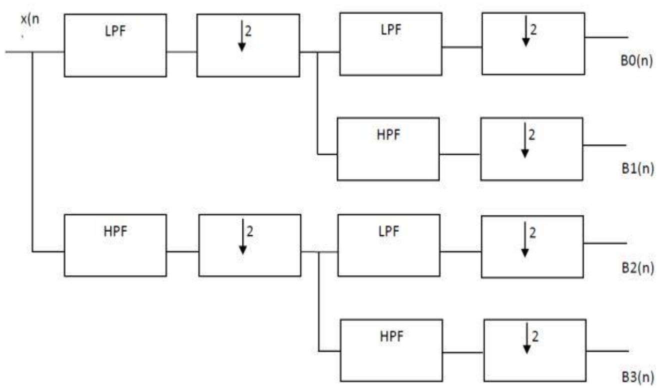


Fig 4. Encoding at transmitter

In the above block diagram the input signal is a speech signal, which is passed through the low pass and

ISSN: 2278 – 909X

high pass filter to split the signal into lower and higher frequency bands. These two signals are down sampled by two in the next step. This down sampled signal by two is further passed through low and high pass filters respectively. Finally 4 signals are down sampled by 2, to get the 4 bands of signal. These four bands of signal are transmitted. Since most of the voice signals are present in the lower frequency bands, bands B2 (n) and B3 (n) will contain less information than compared to B0(n) and B1(n) .

RECEIVER:

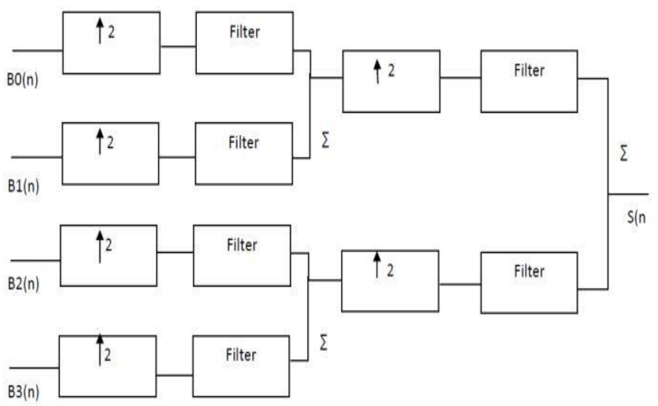


Fig 5.Synthesizing at the receiver end.

Receiver part that is we can call it as the synthesis part also. The input to this block is the encoded signals that are encoded at the transmitter end. These 4 bands signals are up sampled by 2. Then these signals are passed through a low pass filter. In synthesis block low pass filters will act as a smoothing filter. Upper 2 bands lower 2 bands are added to get 2 bands of signals. Now these two signals are further up sampled by two and smoothing is performed by the low pass filter. Outputs from this low pass filter are added to get the final signal, which will resemble the input speech signal that is being processed at the transmitter end.

Since most of the speech energy is contained in the lower frequencies, we would like to encode the lower-frequency band in more bits than the high-frequency band. Sub-band coding is a method where the speech signal is subdivided into several frequency bands and each band is digitally encoded separately with different number of bits.

IV. RESULT AND DISCUSSION:

a) The data rates that are available after executing are specified below:

By specifying the length of the input signal in code as length(x) we founded that the length was 42752 samples. Each band has a length of 10703 samples.

By adding all the 4 bands length we get
10703+10703+10703+10703=42812 samples

Hence the total bandwidth is much more than the input signal. Since the amount of information that is being present in the lower bands is much more than the higher bands as we can see from the above plot. We will assign different bit length to different bands. In this way we can save bandwidth.

If the original signal has been transmitted over the channel, using 16 bit encoding (PCM), the total amount of bits required for the entire data to be transmitted is

42752*32=1368064

Hence the total amount of bits that is being used is 1368064

Now let us assign differ bit size to each band
Band 1=32 bits.
Band 2=24 bits.
Band 3=16 bits.
Band 4=8 bits.

The total amount of bits that is required to encode all bands is being calculated below

(10703*32) + (10703*24) + (10703*16) + (10703*8) =856240

As compared to 1368064, after sub band coding the number of samples are found to be 856240. The number of bits that have been saved is 5, 11,824 samples.

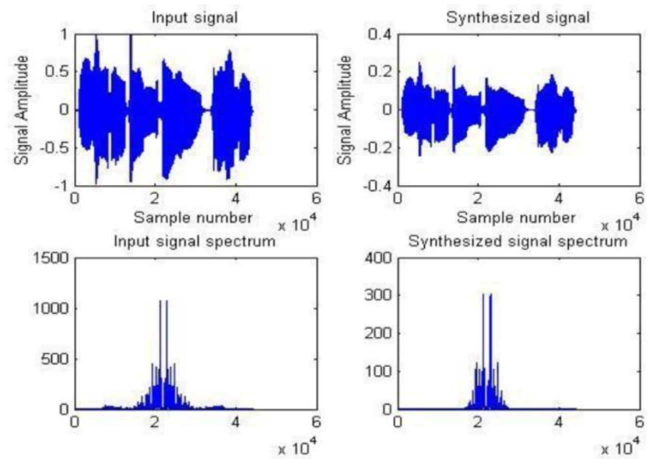
Coding test show that this new sub band speech coding scheme based on multi rate sampling can not only realize the splitting and combining of the speech bands conveniently, but also obtain the high compression ratio coding of speech signal. It provides a flexible variable bit rate speech coding method and suits packet switching network. It allows the switch node to regulate or control the transmission bit rate of speech within a large flexible range actively. Taking correlation tests prove that its performance is satisfying.

V. SIMULATION RESULTS:

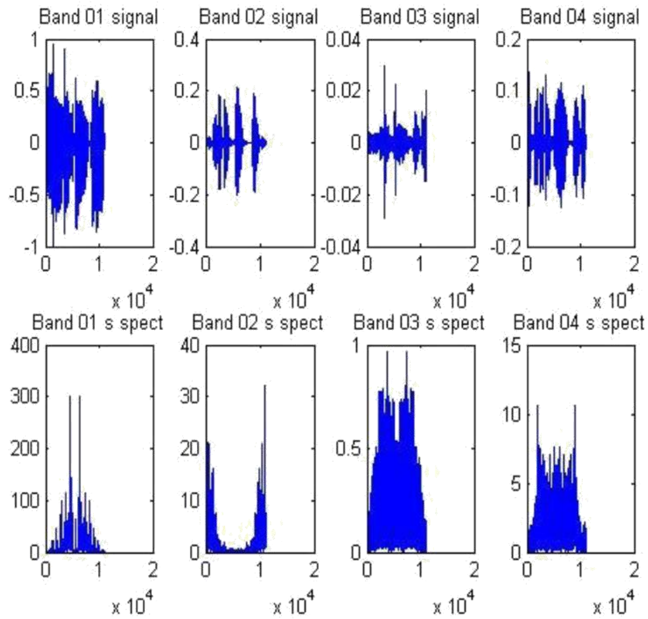
For clean speech signal:

Let the original signal be x1 contain fs samples. Obtained waveforms of every stages are given below.

Cut off frequency Wc = 4 kHz
Order off FIR filter N =20
Input signal x1 = Clean_S.wav



ISSN: 2278 – 909X



Processed signal of 1st low pass filter is represented in frequency domain. We can see that most of the information will lie on this band. The frequency spectrum of band 02, there will be little less information. In band 03 spectrum i.e. in higher frequencies the amount of information that is present is very less and also the info is scattered. Hence only few information signals are present in this band. In the band 04 spectrum, Most of the signal that is present in this band is noise. Amplitude levels of this band are also less.

First bands time domain representation of the signal is shown in the above figure. Since most of the information is present in the lower frequency band, this band almost resembles the original signal. From time domain representation of band 2 signal, we can see that the signal slightly deviates from the original signal. Even band 3 and band 4 times domain signals will have very less information.

VI.CONCLUSION:

Simulation of transmitter and receiver part has been completely carried out successfully and the results obtained are satisfactory. Where the synthesized and input speech were one and the same with very less error in synthesis. The results are verified by taking correlation function for input and synthesized speech Signal. Speech coding is currently an active topic for research in the areas of Very Large Scale Integrated Circuit (VLSI) technologies and Digital Signal Processing (DSP). The proposed paper aims at providing a low bit-rate for telecommunication industries. Lower bit-rate reduces the consumption of bandwidth so that more signals can pass through reducing the cost, power management and complexity in previous algorithms. This paper finds its application in data compression of digital audio signals, mobile telephony and voice over IP.

VII.FUTURE SCOPE:

This paper can be extended to de noising different noise signals like street noise, babble noise, impulse noise etc. This sub band coding concept can be applied to image processing specially can be used for image smoothing.

REFERENCES:

1. Siddhartha Seth, "sub band coding of speech signal using DSP" speech signal processing conference, Stanford University, 2005.

2. M. Ramya and M. Sathyamoorthy, "Speech Coding by using Sub band Coding", International Conference onComputing and Control Engineering (ICCCE 2012), 12 & 13 April, 2012

3. John G. Proakis and Dimitris G. Manolakis, "Digital Signal Processing: Principles, ssAlgorithms and Applications", Third Edition.

4. Ashraf M. Aziz, "Sub band Coding of Speech Signals Using Decimation and Interpolation"13th International Conference on Aerospace Sciences & Aviation Technology, 2009.

5. Arvindraj Desai, Tippanna, Shashidhar Y, Sandeep A.M. VTU Final year project, DayanandSagar College of Engineering, Bengaluru, 2010-2011.

6. SupavitMuangjaroen and ThaweesakYingthawornsuk, "A Study of Noise Reduction in Speech Signal using FIR Filtering", International Conference on Advances in Electrical and Electronics Engineering (ICAEEE'2012) April 13-15, 2012 Pattaya.

7. Martin VONDR´ASˇEK, Petr POLLA´K, "Methods for Speech SNR estimation: Evaluation Tool and Analysis of VAD Dependency", Dept. of Circuit Theory, Czech Technical University, Technick´a, 2009.

8. Zheng-Hua Tan, Senior Member, IEEE, Paul Dalsgaard, Senior Member, IEEE, and Borge Lindberg, Member, IEEE Exploiting, " Temporal Correlation of Speech for Error Robust and Bandwidth Flexible Distributed Speech Recognition", IEEE TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING, VOL. 15, NO. 4, MAY 2007.

ISSN: 2278 – 909X