

# Sub Band Coding of Speech Signal by using Multi-Rate Signal Processing

Vijayakumar Majjagi Student, 3<sup>rd</sup> Semester M.Tech Digital Electronics, G.M.Institute of Technology  
Davangere, Karnataka, India

**Abstract:** Interest in signal processing long predates computers. As long as people have tried to send or receive information through electronic media, such as telegraphs, telephones, television, radar, etc., there has been the realization that these signals may be affected by the system used to acquire, transmit, or process them. Sometimes these systems are imperfect and introduce noise, distortion, or other artifacts. Understanding the effects these systems have and finding ways to correct them is the foundation of signal processing. There are many types of signal processing. Among that Digital signal processing is more efficient and widely used. Multirate systems are building blocks commonly used in digital signal processing (DSP).

In conventional speech processing applications, speech signal is encoded using fixed number of bits over the entire speech signal band. During the process, the bandwidth requirement for speech transmission is relatively high which is of concern. The QMF (Quadrature Mirror Filter) banks are the fundamental building blocks for spectral splitting. The technique is developed to design the so-called perfect reconstruction QMF bank, which allows complete elimination of amplitude and phase distortion of the reconstructed signal. The low pass filtered signal is decimated and encoded with more number of bits and high pass filtered signal is also decimated and encoded with less number of bits. These two bit streams are multiplexed and transmitted. In receiver side the received signal is de-multiplexed and decoded. The signal is passed through the interpolators and then through the synthesis filters so as to reconstruct the speech signal. The reconstructed signal is compared with the original speech signal.

## I. Introduction

A multirate DSP system simply uses more than one sampling rate within the system. In many systems, multirate DSP increases processing efficiency, which reduces DSP hardware requirements. Also, a few systems are inherently multirate, for example, a "sampling rate converter" system that converts an input sampling rate to a different output sampling rate. Multirate systems play a central role in many areas of signal processing, such as filter bank theory and multiresolution theory, they are essential in various standard signal-processing techniques such as signal analysis, denoising, and compression and so on. During the last decade, however, they have increasingly found applications in new and emerging areas of signal processing, as well as in digital communications.

"Multirate" means "multiple sampling rates". A multirate DSP system uses multiple sampling rates within the system. Whenever a signal at one rate has to be used by a system that expects a different rate, the rate has to be

increased or decreased, and some processing is required to do so. Therefore "Multirate DSP" refers to the art or science of changing sampling rates. "Resampling" means combining interpolation and decimation to change the sampling rate by a rational factor. Resampling is done to interface two systems with different sampling rates.

Multirate DSP consists of:

1. **Decimation:** It is a process to decrease the sampling rate.
2. **Interpolation:** It is a process to increase the sampling rate.

"Down sampling" is a process of removing some samples, without the low pass filtering. A signal is down sampled only when it is "oversampled"(i.e. sampling rate > Nyquist rate). This combined operation of filtering and down sampling is called "Decimation". To down sample by a factor of M, we must keep every M<sup>th</sup> sample as it is and remove the (M-1) samples in between.



Fig 1.1: Symbol of down sampler

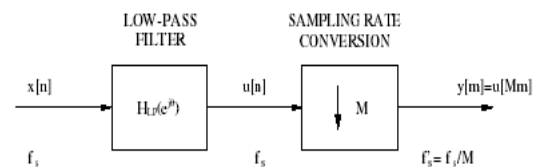


Fig 1.2: Block diagram of a decimator

"Up sampling" is the process of inserting zero-valued samples between original samples to increase the sampling rate. (This is called "zero-stuffing"). Given a sequence  $x[n]$ , we can define

$$x_u[n] = \begin{cases} x[n/L], & n = 0, \pm L, \pm 2L, \dots \\ 0, & \text{otherwise} \end{cases}$$

Where  $x_u[n]$  is the sequence up-sampled from  $x[n]$  by a factor of L. This means that  $x_u[n]$  is generated by padding (L-1) zeros between every sample of  $x[n]$ .

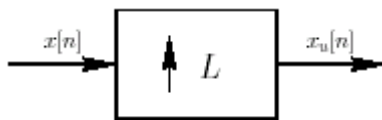


Fig 1.3: Symbol for up-sampler

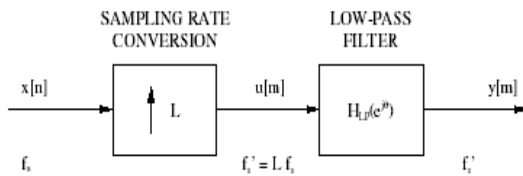


Fig 1.4: Block diagram of an interpolator

"**Interpolation**" is the process of upsampling followed by filtering (to remove the undesired spectral images.) The result is a signal sampled at a higher rate. The interpolation factor ( $L$ ) is the ratio of the output rate to the input rate.

## II. Basics of Speech Processing

Speech is the most basic and preferred means of communication amongst humans. Even though, one can communicate information textually using a teletype, at almost the same rate as a person speaking the same text, but spoken message communication is preferred, as it carries much more information like speaker's identity, emotional state and prosodic nuances which add to naturalness in communication. Hence, there is insatiable demand for voice communication. Digital cellular and satellite telephony, tele-conferencing, voice messaging, voice communication over internet telephony are just a few of prominent everyday modern applications that are driving this demand.

Most of these incorporate mechanisms to: provide speech waveform matching, represent the spectral properties of speech, and to optimize the coder's performance for the human ears. The coding technology is being spurred by advances in several fields— better modeling of human speech production and perception system, simultaneous evolution of device technology to support substantial amount of real-time digital signal processing and storage of digital data.

## III. 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 LPP transfer function  $H_a(S)$  meeting these specifications and then to transform it into the desired digital filter transfer function  $H(Z)$ . This approach has been widely used for many reasons,

1. Analog approximation techniques are highly advanced.
2. They usually yield closed form solutions.
3. Extensive tables are available for analog filter design.

The basic idea behind the conversion of an analog prototype transfer function  $H_a(s)$  is to apply a mapping from the  $S$ -domain to the  $Z$ -domain so that the essential properties of the analog frequency response are preserved.

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

$$H(Z) = \sum_{n=0}^{N-1} h[n]z^{-n}$$

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

$$h(N - 1 - n) = h[n] \text{ or } -h[n]$$

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

## IV. Structure of FIR filter

Now that we have seen how the parts make a filter, we will demonstrate some FIR filters and discuss some important characteristics: 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 ( $b_0, b_1, \dots, b_K$ ) 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. With the structure of Figure 2a.2, it is possible to determine the output. It is also possible to determine an equation for the output, which is

$$y[n] = b[0]x[n - 0] + b[1]x[n - 1] + \dots + b[K]x[n - K]$$

Notice that whatever index is used for  $b[.]$  is also used in  $x[n - .]$ . This means we can represent everything on the right hand side of the equation as a summation.

$$y[n] = \sum_{k=0}^K b[k]x[n - k]$$

## V. Sub band Coding

Sub Band Coding (SBC) is a frequency domain coding technique in which the input signal is decomposed into a number of sub bands so that each of these frequency bands can be encoded separately.

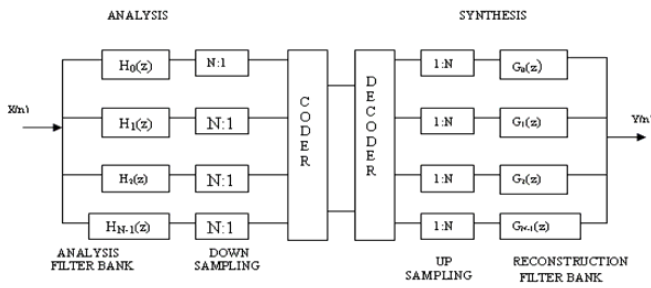


Fig 5.1: Block Diagram of Sub-Band Coding

Sub-Band Coding (SBC) is a powerful and general method of encoding audio signals efficiently. Unlike source specific methods (like LPC, which works only on speech), SBC can encode any audio signal from any source, making it ideal for music recording, movie soundtrack. MPEG Audio is the most popular example of SBC.

The basic idea behind SBC is 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 sense lower energy at nearby frequencies. We say that the louder frequency *masks* the softer frequencies. The louder frequency is called the *masker*. Strictly speaking, what we're describing here is really called *simultaneous* masking (masking across frequency). There are also non-simultaneous masking (masking across time) phenomena, as well as many other phenomena of human hearing, which we're not concerned with here.

The basic idea of SBC is to save signal bandwidth 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 make out the difference.

A variety of techniques have been developed to efficiently represent speech signals in digital form for either transmission or storage. 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.

In the sub band-coding system the input signal, after being sampled at its Nyquist rate, is divided into channels by first being passed through a bank of low pass and high pass filters. The output of each filter is decimated to a rate determined by the number of sub bands and then each of these channel outputs are encoded separately. At the receiver the signals, after being decoded, are interpolated back to the original sampling rate by a bank of interpolation filters and then are summed to reconstruct the input signal. It is important that in subband coding systems the individual channel signals be decimated in such a way that the number of samples coded and transmitted does not exceed the number of samples in the original signal since this number is necessary and sufficient for the recovery of the original signal.

Transmitter

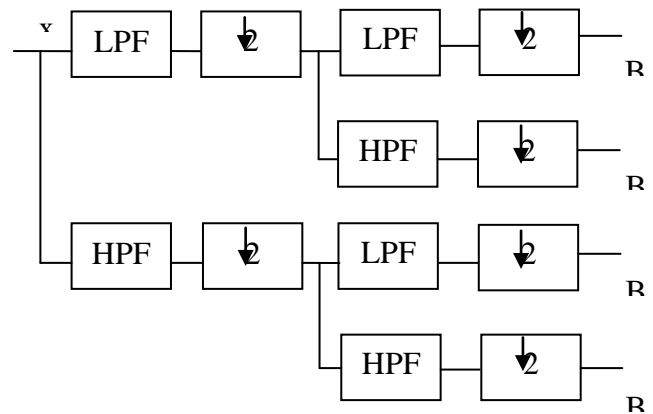


Fig 5.2: Encoding at transmitter

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

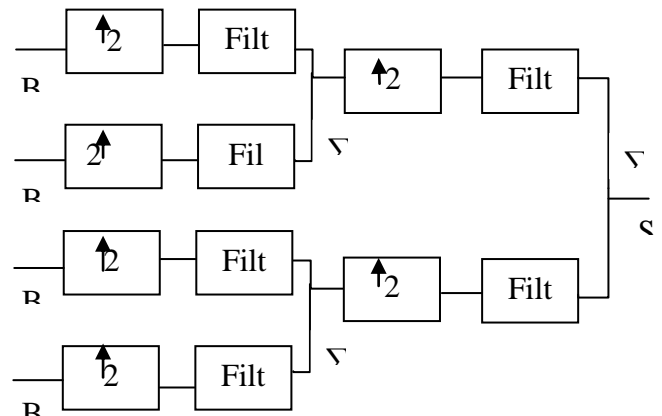


Fig 5.3: Synthesizing at the receiver end.

Receiver part that is we can call it as the synthesis part also. The inputs to this block are 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.

## VI. Two channel QMF bank

In many applications, a discrete-time signal  $x[n]$  is first split into a number of sub band signals by means of an analysis filter bank; the sub band signals are the sub band signals are then processed and finally combined by a synthesis filter bank resulting in an output signal  $y[n]$ . If the sub band signals are band limited to frequency ranges much smaller than that of the original input signal, they can be downsampled before processing. Because of the lower sampling rate, the processing of the down-sampled signals can be carried out efficiently. After processing, these signals are upsampled before being combined by the synthesis bank into a higher-rate signal. The combined structure employed is called a Quadrature-mirror filter (QMF) bank. If the down-sampling and up-sampling factors are equal to or greater than the number of bands of the filter bank, then the output  $y[n]$  can be made retain some or all of the characteristics of the input

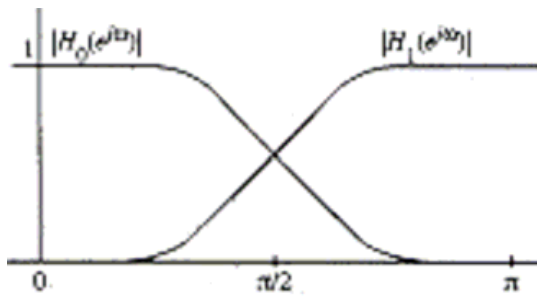


Figure 6.1 Frequency Response Characteristics of QMF Bank

$x[n]$  by properly choosing the filters in the structure. The two channel Quadrature Mirror Filter (QMF) bank is multirate digital filter structure that employs two down-samplers in the “signal analysis” section and two upsamplers in the “signal synthesis” section. The input signal  $x[n]$  is first passed through a two-band analysis filter bank containing the low pass and high pass filters with frequency responses  $H_0(z)$  and  $H_1(z)$ . Their corresponding impulse responses are  $h_0(n)$  and  $h_1(n)$  respectively, with a cutoff frequency at  $\pi/2$ , as shown in the fig. The frequency response characteristics of QMF bank. The sub-band signals  $V_0(n)$  and  $V_1(n)$  are then down-sampled by a factor of 2.

Each down-sampled sub band signal is encoded by exploiting the special spectral properties of the signal, such as energy levels and perceptual importance. The coded sub-band signals are combined into one sequence by multiplexing and

either stored for later retrieval or transmitted. At the receiving end, the coded sub-band signals are first recovered by demultiplexing and decoders are used to produce approximations of the original down-sampled signals. The decoded signals are then up-sampled by a factor of 2 and passed through the synthesis filter bank composed of the low pass and high pass filters whose frequency responses are  $F_0(z)$  and  $F_1(z)$  whose outputs are then added yielding  $y[n]$ . It follows from the figure that the sampling rates of the input signal  $x[n]$  and output signal  $y[n]$  are the same. The analysis and the synthesis filters in the QMF bank are chosen so as to ensure that the reconstructed output  $y[n]$  is a reasonable replica of the input  $x[n]$ .

## VII. Results and Conclusion

We have successfully implemented the sub-band coding system by designing an optimum four channel QMF bank. The frequency response characteristics of LPF and HPF used in QMF bank are as given in fig 7.1: From the above characteristics it is seen that, the response of the QMF filter is almost approaching the ideal all-pass filter characteristic, which results in perfect reconstruction of the input speech signal.

### Input speech signal and its specifications

The speech signal on which sub-band coding is to be performed is given as an input to the QMF bank, which was discussed in the previous chapters. For this we recorded a speech signal using the tool sound recorder i.e available in the windows with the following specifications

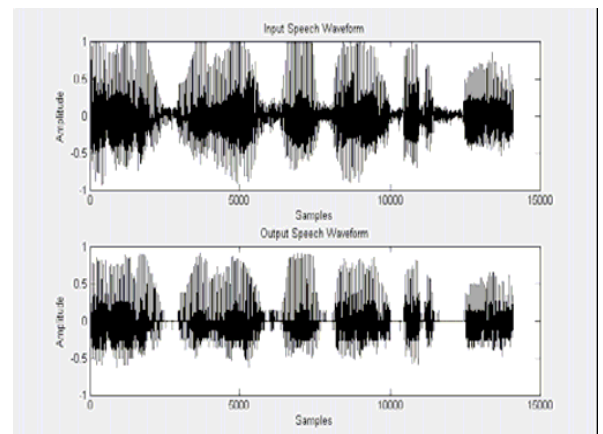


Fig 7.1: Frequency Response of QMF Bank

The recorded speech signal is of two seconds duration with a length of 21600 samples. The speech signal is sampled with a sampling frequency of 8 kHz and coded with 8 bits per sample. The input speech waveforms and output reconstructed speech waveforms are shown below:

From these waveforms we can observe that there is a delay of 31 samples between the input and output speech, which is equal to  $N-1$  (where  $N=32$  is the length of the filter).

The data rate reduction depends upon the number of bits allocated for low-pass and high-pass sections.

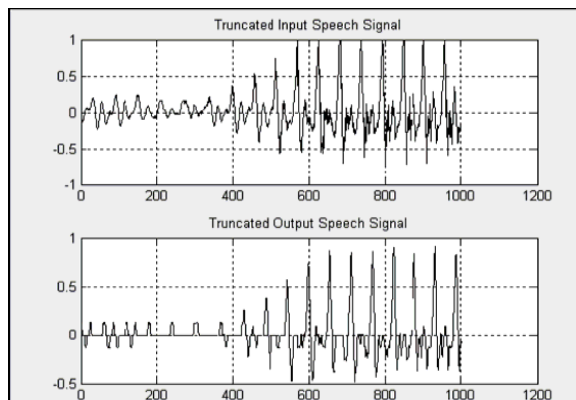


Fig 7.2: Truncated input and output speech signal

Format: PCM

Attributes: 8 kHz, 8 bit, Mono

### VIII. Conclusion

The intention of this work is to design and implement a SUBBAND CODING system. We have successfully designed an optimum low pass filter for four channel QMF Bank to minimize the amplitude distortion. From this Low pass filter we have designed a High pass Filter. Using these filters we have successfully simulated a two channel QMF bank for sub-band coding of input speech signal. The result shows that the output is a perfect reconstruction of the input speech signal.

### ACKNOWLEDGMENT

It is a pleasure to recognize the many individual who have helped me in completing this technical paper.

Mrs. Nethravathi U.M (G.M.I.T Davangere) for all the technical guidance, encouragement and analysis of the data throughout this process.

### REFERENCES

- [1] "Digital Signal Processing (Principles, Algorithms and Applications)" by John G.Proakis and Dimitris G.Manolakis.
- [2] Digital Signal Processing, A. Oppenheim & R. Schaffer, (Prentice-Hall, 1975, ISBN 0-13-214635- 5).
- [3] P.P. Vaidyanathan. Multirate Systems and Signal processing. Prentice-Hall, Englewood Cliffs, NJ, 1993.
- [4] S.K. Mitra. Digital Signal Processing – A Computer-Based Approach. Mc Graw-Hill, New York, 2 edition, 2001.
- [5] [Chi, et al.] Chi, T., Gao, Y., Guyton, M., Ru, P., and Shamma S.A. "Spectro-Temporal Modulation Transfer Functions and Speech Intelligibility".