

# MULTI-RATE SIGNAL PROCESSING SUB-BAND CODING FOR SPEECH SIGNAL

Dr.I.Kullayamma<sup>1</sup> & Mr.M.Venkataresh<sup>2</sup>

**Abstract**— In many applications of digital signal processing such as Tele-communication systems, transmitting different types of signals with different rates of Band widths and retaining the quality of speech signal is an important criteria. An Effective utilization of the Band width and proper sampling rate conversion enables usage of channel capacity productively. In this paper we have studied several characteristics of multi-rate Digital signal processing systems using FIR,IIR filters and implemented a frequency domain Digital coding technique, sub-band coding on a sinusoidal, real time speech signal in an audible frequency spectrum of 20-20KHz and 22100 samples. The initial stage involves in sub-dividing the input signal into several frequency bands and each frequency band is digitally encoded separately with different no. of bits and at the receiver end the stream of bits are decoded and a small delay is observed in between input and reconstructed signal. several metrics like Variances ,Arithmetic, Geometric Mean, Coding gain are evaluated on the input signal.

**Index Terms:** Multi-rate Digital signal processing, sub-band coding, FIR and IIR filters and Coding gain

## 1. INTRODUCTION

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"[1] refers to the art or science of changing sampling. This paper give study of different muti-rate signal proccessing techniques and implementation of a frequency domain multi-rate technique i.e. sub-Band coding on an real time speech signal. The important elements of a multi-rate signal processing systems are Up-sampler - Used to increase the sampling rate by an integer factor, Down-sampler - Used to decrease the sampling rate by an integer factor.

Sub Band Coding (SBC) 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. Sub-Band Coding 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.

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 quantized 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[1]. It is important that in sub-band coding systems the individual channel signals be decimated [2] 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.

## 2. CLASSIFICATIONS OF MULTI RATE DSP SYSTEMS

### 2.1 Up sampler

An up-sampler with an up-sampling factor L, where L is a positive integer, develops an output sequence  $x_u[n]$  with a sampling rate that is L times larger than that of the input sequence  $x[n]$ [2]. Block-diagram representation of up-sampler is as shown in fig.1 also called a sampling rate expander or simply an expander.



Fig 1: Block Diagram of Up-sampler

Up-sampling operation is implemented by inserting (L-1) equidistant zero-valued samples between two consecutive samples of input sequence  $x[n]$ [2].

<sup>1</sup> Asst.Professor ,Dept.Of ECE, S.V.U College of Engineering, S.V University

<sup>2</sup> Asst.Professor ,Dept.Of ECE, Sree Vidyanikethan Engineering College.

Input-output relation is:

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

In terms of the z-transform, the input-output relation is then given by

$$\begin{aligned} X_u(z) &= \sum_{n=-\infty}^{\infty} x_u(n)z^{-n} = \sum_{\substack{n=-\infty \\ n \text{ even}}}^{\infty} x(n/2)z^{-n} \\ &= \sum_{m=-\infty}^{\infty} x[m]z^{-2m} = X[z^2] \end{aligned}$$

### 2.2 Down Sampler

A down-sampler with a down-sampling factor  $M$ , where  $M$  is a positive integer, develops an output sequence  $Y[n]$  with a sampling rate that is  $1/M$  th of that of the input sequence  $x[n]$ .

Block-diagram representation of down-sampler is shown in figure 2 and is also known as sampling rate compressor.



Fig 2: Block Diagram of Down sampler

Down-sampling operation is implemented by keeping every  $M$ -th sample of  $x[n]$  and removing  $(M-1)$  [2] in-between samples to generate the output sequence  $y[n]$ .

Input-output relation is:

$$y[n] = x[Mn]$$

Applying the z-transform to the input-output relation of a factor-of- $M$  down-sampler

$$Y(z) = \sum_{n=-\infty}^{\infty} x[Mn]z^{-n}$$

The desired input-output relation in the transform domain for down-sampler is:

$$Y(Z) = \frac{1}{M} \sum_{k=0}^{M-1} X(Z^{1/M} W_M^{-k})$$

### 2.3 Decimator

So to avoid aliasing first we restrict the bandwidth of base band signal to  $\pi/M$  by passing the signal through a lowpass filter of pass band equal to  $\pi/M$

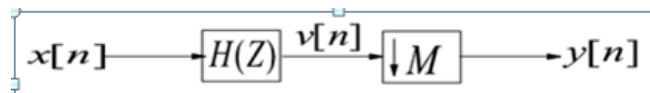


Fig.3 Block Diagram of Decimator

The above system is called a decimator and the low pass filter  $H(z)$  is called decimation filter band limits the input signal prior to downsampling[3] to ensure no aliasing.

As said, to insert images in Word, position the cursor at the insertion point and either use Insert | Picture | From File or copy the image to the Windows clipboard and then Edit | Paste Special | Picture (with "Float over text" unchecked).

The authors of the accepted manuscripts will be given a copyright form and the form should accompany your final submission.

### 3. DIGITAL FILTERS

#### 3.1 Fir And Iir Filters

Digital Filters are classified depending on the duration of their unit impulse response or by their structure. When a filter produces a unit impulse response that has an infinite duration it is called an Infinite Impulse Response (IIR) filter. If the digital filter has a unit impulse response having a finite duration then it is Finite Impulse Response (FIR) filter.

According to the structure, if the output of the filter  $y[n]$  is a function of future, [4] current and past inputs and as well as past outputs such a filter is known as recursive filter.

$$y[n] = f(\dots x[n-1], x[n], x[n+1], \dots, y[n-1] \dots)$$

If the output of a filter is a function of only present, past and future inputs such a filter is known as non-recursive filter.

$$y[n] = f(\dots x[n-1], x[n], x[n+1] \dots)$$

General forms of IIR and FIR filters are:

$$H(Z) = \frac{\sum_{k=0}^M a_k z^{-k}}{1 + \sum_{k=1}^N b_k z^{-k}}$$

$$H(Z) = \sum_{k=0}^M a_k z^{-k}$$

Differences of IIR and FIR Filters are as follows

- IIR filters are characterized by rational system function whereas FIR system is characterized by irrational system function.
- IIR filters are of infinite duration and does not have linear phase and these are used where some phase distortion is tolerable, FIR filters are of finite duration have linear phase.
- IIR filters are unstable and are realized by recursive method, FIR filters are stable and are realized by non-recursive method.
- IIR filters require less number of arithmetic operations and they have low computational complexity and smaller memory requirements, higher order FIR filters have sharp amplitude response.
- FIR filters are linear and stable taking this in to consideration, we considered FIR filter in our consideration

To implement a digital filter on a digital computer the input and output relation must be converted to a computational algorithm. The algorithm must be specified in terms of a set of basic computations or elements. These basic operations are addition, delay and multiplication by a constant. There are four realization techniques and those are:

- Direct form-I realization
- Direct form-II realization
- Cascade realization
- Parallel realization

Above four realization techniques require more number of adders and delay elements. The computational complexity, adders and delay elements are reduced by realizing the filters using poly-phase decomposition.

#### 3.2. Nyquist Filters

Nyquist filters are a special class of filters which are useful for multi-rate implementations. Nyquist filters also find applications in digital communications. Systems where they are used for pulse shaping[5] (often simultaneously performing multirate duties). Nyquist filters are typically designed via FIR approximations. However, IIR designs are also possible. The cascade of Nyquist interpolators (or decimators) results in an overall Nyquist interpolator (or decimator). This makes Nyquist filters highly desirable for multistage designs.

Under certain conditions, a low pass filter can be designed to have a number of zero-valued coefficients. When used as interpolation filters, these filters preserve only the non-zero samples of the up-sampler output at the interpolator output. Moreover, the presence of these zero-valued coefficients, these filters are computationally more efficient than other lowpass filters of same order. These filters are often used in single-rate and multi-rate signal processing. Consider the factor-of- $L$  interpolator shown in Fig 3.1

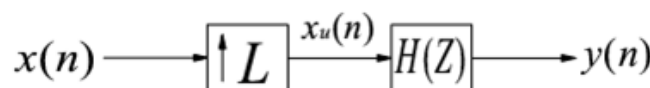


Fig 4: Block Diagram of Interpolator

The input-output relation of interpolator in Z-domain is given by

$$Y(Z) = H(Z)X(Z^L)$$

If  $H(Z)$  is realized in the  $L$ -band polyphase form, then we have

$$H(Z) = \sum_{i=0}^{L-1} Z^{-i} E_i(Z^L)$$

Assume that the  $K$ -th poly-phase component of  $H(Z)$  is a constant, i.e.,  $E_K(Z) = \alpha$

$$H(Z) = E_0(Z^L) + Z^{-1}E_1(Z^L) + \dots + Z^{-(k-1)}E_{k-1}(Z^L) + \alpha Z^{-k} + Z^{-(k+1)}E_{k+1}(Z^L) + \dots + Z^{-(L-1)}E_{L-1}(Z^L)$$

Then we can express  $Y(Z)$  as,

$$Y(z) = \alpha Z^{-k} X(Z^L) + \sum_{L=0, L \neq K}^{L-1} Z^{-1} E(Z^L) X(Z^L)$$

As a result,

$$y[Ln + K] = \alpha x[n]$$

Thus the input samples appears at the output without any distortion for all values of  $n$ , whereas in between  $(L - 1)$  samples are determined by interpolation. The filter with above property is called Nyquist filter and Nyquist filter is known as half-band filter

#### 4. QUADRATURE MIRROR FILTER & SUB BAND CODER

The QMF (Quadrature Mirror Filter) [6] banks are the fundamental building blocks for spectral splitting. The QMF structure allows spectral decomposition into continuous overlapping subbands in such a way that aliasing incurred in the initial “analysis” stage is eliminated during signal reconstruction by the “synthesis” stage. A QMF bank can be formed from FIR or IIR filters.

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 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 down sampled before processing. Because of the lower sampling rate, the processing of the down-sampled signals can be carried out efficiently.

The Frequency Response Characteristics of 2-channel QMF Bank [6] is shown in fig 5 the analysis filter bank is designed using FIR filters.

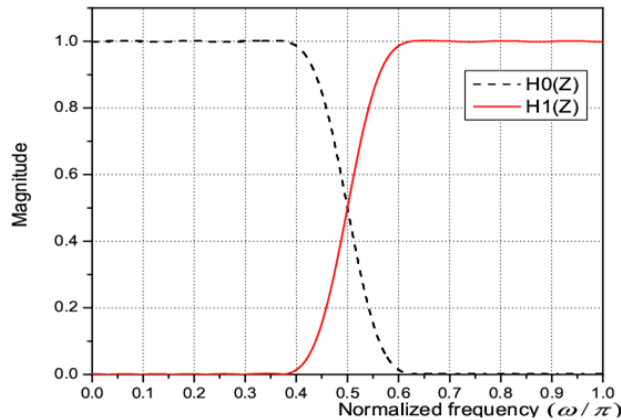


Fig5: Frequency response of QMF Filter for N=2

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[7]. 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. SBC is used 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.

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. So in Sub-band coding 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[8] 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 quantized 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 sub-band 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.

The coding gain of the system is increased by quantizing the sub-band signals. The block diagram of sub-band coding system is as shown in fig.6

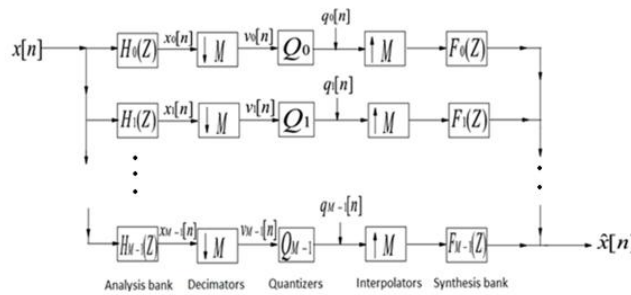


Fig 6: M Channel Sub-Band Coder

Coding gain of the system is given by the equation:

$$G_{SBC}(M) = \frac{\text{arithmetic mean of variances of subband signals}}{\text{geometric mean of variances of subband signals}}$$

So  $G_{SBC}(M)$  can be written as:

$$G_{SBC}(M) = \frac{\frac{1}{M} \sum_{k=0}^{M-1} \sigma_{x_k}^2}{(\prod_{k=0}^{M-1} \sigma_{x_k}^2)^{1/M}}$$

**5. RESULTS AND DISCUSSIONS**

We simulated the sub-band coder using two signals one is sinusoidal signal and other signal is real time speech signal.

The input sinusoidal form signal  $X[n] = \cos(2\pi(0.4)n) + 0.5\cos(2\pi(0.6)n)$  of length 120 is shown in the figure 7.

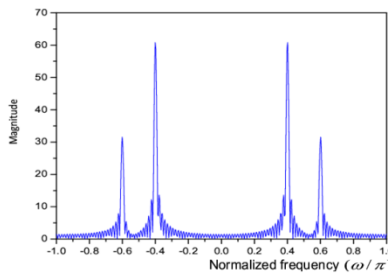


Fig.7: Input sinusoidal Signal

we taken the sinusoidal signal of frequencies and the spectrum of the same is as shown in above fig. this signal is passed through analysis bank and then decimated and each sub-band signal is quantized using 8 quantization levels and the output of the stage is as shown in figure 8

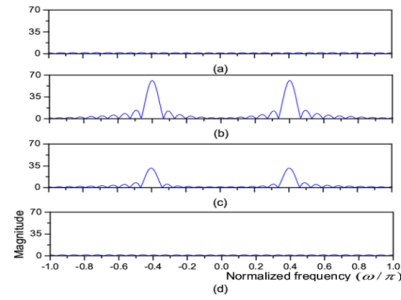


Fig 8: (a) quantized sub-band signal one (b) quantized sub-band signal two (c) Quantized sub-band signal three (d) quantized sub-band signal four And these signals are transmitted and interpolated by 4 and filtered at the receiver and sub-band signals are added to recover the message signal.

An input speech signal of 22100 samples in time domain as shown in figure 9 is taken . This speech signal is passed through the analysis filter bank.

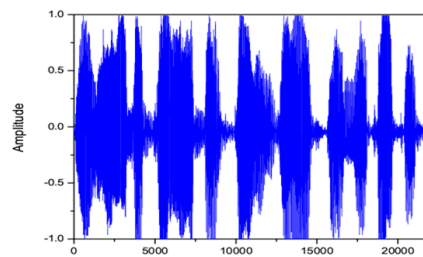


Fig 9: Input speech signal of 22100 samples

As we know that information is at low frequencies more information resides in subband 1. And each subband signal is decimated by 4 as we are considering 4 subbands and then quantized.

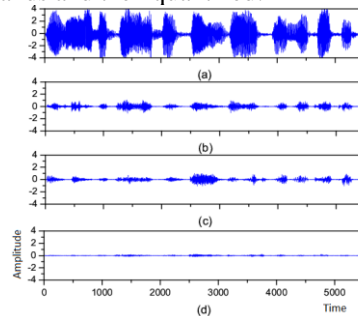


Fig 10: (a) quantized sub-band signal one (b) quantized sub-band signal two (c) Quantized sub-band signal three (d) quantized sub-band signal four

As we decimate each sub-band signal by 4, 22100 samples are reduced to 5525 samples that is  $(22100/4)$ .

These quantized sub-band signals are encoded using different number of bits for different sub-bands for efficient transmission and then transmitted. At the receiver side first we decode the received sub-band signals using same number of bits used in the encoder at transmitter side and then interpolated by 4. And the interpolated quantized sub-bands are as shown in figure 11.

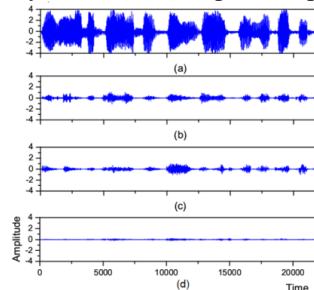


Fig 11: (a) upsampled sub-band signal one (b) up-sampled sub-band signal two (c) up-sampled sub-band signal three (d) up-sampled sub-band signal

As the sub-bands signals are up-sampled by 4, 5525 samples are increased to 22100 samples. These interpolated sub-bands are passed through synthesis filter bank here we had taken length 512 FFT for designing FIR filters so when In many practical applications of digital signal processing, such as telecommunication systems, one is faced with the problem of changing the sampling rate of a signal, either increasing it or decreasing it by some amount. the filtered sub-band signals are added to reconstruct the original signal there is delay of 512 samples in the recovered signal but these 512 samples delay at output is tolerable because the original signal has 22100 samples. Human ear can't identify this small delay. There is no amplitude and phase distortions in the reconstructed signal.

The recovered and original speech signals are as shown in fig 12.

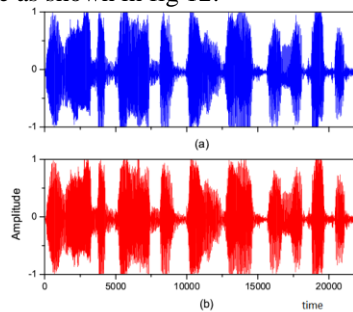


Fig 12: (a) original speech signal (b) recovered speech signal

The metrics like Arithmetic mean(AM) , Geometric mean(GM) ,Coding gain(CG), for both sinusoidal and speech signal are evaluated and showed in Table I

Table I: Comparison of AM,GM and CG for Sinusoidal and Real time Speech Signal

S.NO	Sinusoidal Signal			Speech Signal of 22100 Samples		
	AM	GM	CG(dB)	AM	GM	CG(dB)
1.	0.12	0.20	5.93	0.03	0.003	10.43

## 6. CONCLUSION

The time domain and frequency domain characteristics of various Multi rate digital signal processing system are studied and a frequency domain Multi-rate DSP technique, Sub-Band Coding is implemented on a Testimonial sinusoidal signal and a real time speech signal of 22100 samples and metrics like arithmetic mean , geometric mean and coding gain are evaluated for both sinusoidal and real time speech signals and significant improvement is observed in all the metrics

## 7. REFERENCES

- [1] John G. Proakis and Dimitris G. Manolakis. Digital Signal Processing, Principles, Algorithms, and Applications. Prentice Hall. New Jersey, 2008.
- [2] Roberts R. A. and Mullis C. T. Digital Signal Processing. Addison-Wesley, Reading, Mass, 2006.
- [3] Oppenheim A. V. and Schafer R. W. Discrete-Time Signal Processing. Prentice Hall, Englewood Cliffs, New Jersey, 2007.
- [4] Crochiere R. E. and Rabiner L. R. Multirate Digital Signal Processing. Prentice Hall, Englewood Cliffs, New Jersey, 1983.
- [5] Schafer R. W. and Rabiner L. R., "A Digital Signal Processing Approach to Interpolation," Proc. IEEE, Vol. 61, pp. 692-702, June 2003.
- [6] McGillem C. D. and Cooper G. R. Continuous and Discrete Signal and System Analysis, 2nd ed., Holt Rinehart and Winston, New York, 1984.
- [7] Crochiere R. E. and Rabiner L. R., "Optimum FIR Digital Filter Implementations for Decimations, Interpolation, and Narrowband Filtering," IEEE Trans. On Acoustics, Speech, and Signal Processing," Vol. ASSP-23, pp. 444-456, Oct. 2004.
- [8] Crochiere R. E. and Rabiner L. R., "Further Considerations in the Design of Decimators and Interpolators," IEEE Trans. on Acoustics, Speech, and Signal Processing," Vol. ASSP-24, pp. 296-311, August 2007.
- [9] Crochiere R. E. and Rabiner L. R., "Interpolation and Decimation of Digital Signals – A Tutorial Review," Proc. IEEE, Vol. 69, pp. 300-331, March 2008.