www.Vidyarthiplus.com

UNIT - V

APPLICATIONS

MULTIRATE SIGNAL PROCESSING

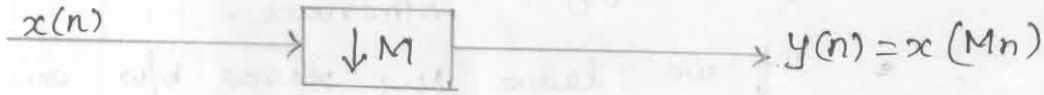
The discrete-time systems that process data at more than one sampling rate are known as multirate systems.

There are many cases where multirate signal processing is used. They are

1. In high quality data acquisition and storage systems.
2. In speech processing to reduce the storage space or the transmitting rate of speech data.
3. In transmultiplexers.
4. Narrowband filtering for fetal ECG and EEG.

DOWNSAMPLING

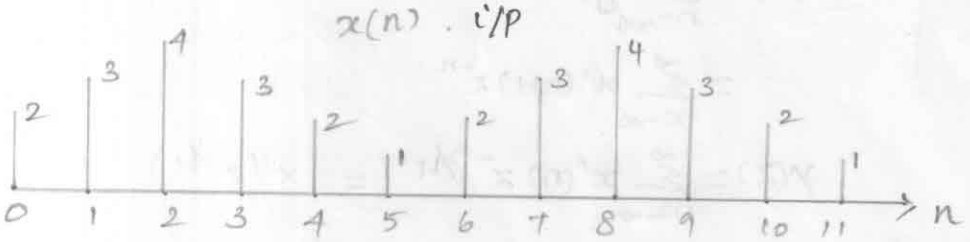
The sampling rate of a discrete-time signal $x(n)$ can be reduced by a factor M by taking every M-th value of the signal.

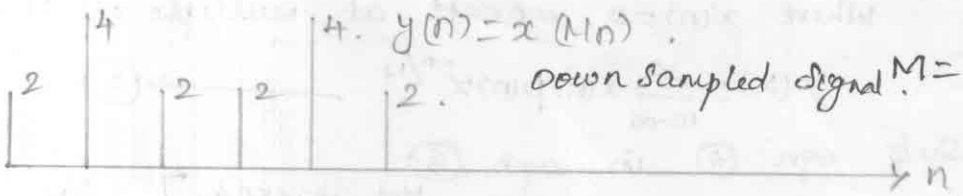


A DOWNSAMPLER

The output signal $y(n)$ is a downsampled signal of the input signal $x(n)$ and can be represented by

$$y(n) = x(Mn)$$





(Left (M-1) samp in b/w x(n) to generate y(n)

DOWNSAMPLING

www.Vidyarthiplus.com

www.Vidyarthiplus.com

SPECTRUM OF THE DOWNSAMPLED SIGNAL.

Let us consider the i/p signal $x(n)$. The z-transform of the signal $x(n)$ is given by

$$X(z) = \sum_{n=-\infty}^{\infty} x(n) z^{-n} \longrightarrow (1)$$

The downsampled signal $y(n)$ is obtained by multiplying the sequence $x(n)$ with a periodic train of impulses $p(n)$ with a period M and then leaving out the M-1 zeros b/w each sample. The periodic train of impulses is given by

$$P(n) = \begin{cases} 1 & n = 0, \pm M, \pm 2M \\ 0 & \text{otherwise.} \end{cases}$$

The discrete Fourier series representation of the signal $p(n)$ is given by

$$P(n) = \frac{1}{M} \sum_{k=0}^{M-1} e^{j2\pi kn/M}. \longrightarrow (2)$$

$x'(n) = x(n)\, p(n)$. ie>

$$x'(n) = \begin{cases} x(n) & n = 0, \pm M, \pm 2M \\ 0 & \text{otherwise.} \end{cases}$$

If we leave M-1 zeros b/w each sample, we get

$$y(n) = x'(nM) = x(nM)\, P(nM)$$
$$= x(nM).$$

The z-transform of the o/p sequence is given by

$$Y(z) = \sum_{n=-\infty}^{\infty} y(n) z^{-n}$$
$$= \sum_{n=-\infty}^{\infty} x'(nM) z^{-n}$$
$$Y(z) = \sum_{n=-\infty}^{\infty} x'(n) z^{-n/M} = X'(z^{1/M})$$

where $x'(n) = 0$ except at multiple of M.

$$Y(z) = \sum_{n=-\infty}^{\infty} x(n) p(n) z^{-n/M} \longrightarrow (3)$$

Sub eqn (2) in eqn (3)

$$Y(z) = \sum_{n=-\infty}^{\infty} x(n) \left[\frac{1}{M} \sum_{k=0}^{M-1} e^{j2\pi kn/M} \right] z^{-n/M}$$

www.Vidyarthiplus.com

www.Vidyarthiplus.com

$$= \frac{1}{M} \sum_{k=0}^{M-1} \sum_{n=-\infty}^{\infty} x(n) \left(e^{-j2\pi k/M} z^{1/M}\right)^{-n}$$

$$Y(z) = \frac{1}{M} \sum_{k=0}^{M-1} X\left(e^{-j2\pi k/M} z^{1/M}\right) \longrightarrow (4)$$

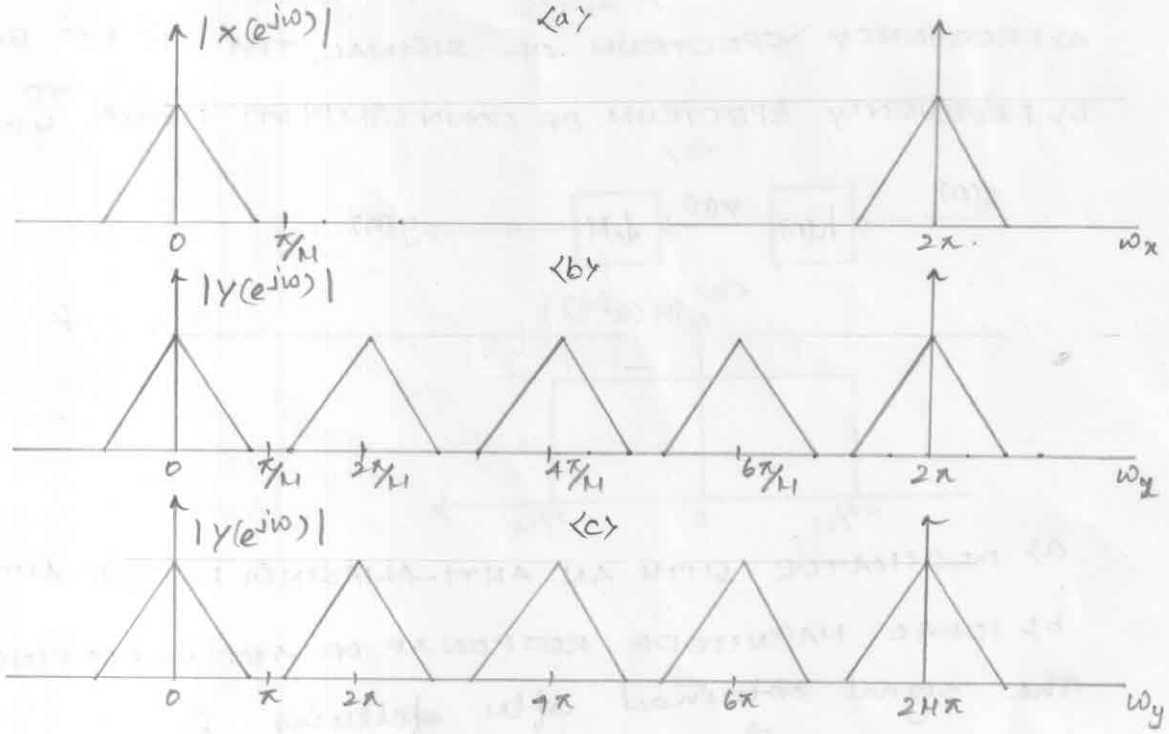
substituting $z = e^{j\omega}$ in eqn (4)

$$Y(e^{j\omega}) = \frac{1}{M} \sum_{k=0}^{M-1} X\left(e^{-j2\pi k/M} e^{j\omega/M}\right)$$

$$Y(e^{j\omega}) = \frac{1}{M} \sum_{k=0}^{M-1} X\left(e^{j(\omega - 2\pi k)/M}\right)$$

$X(e^{j\omega}) \rightarrow$ Fourier Transform of the i/p signal of a downsampler

$Y(e^{j\omega})$ is the F.T of the o/p signal $y(n)$ is a sum of M uniformly shifted and stretched versions of $X(e^{j\omega})$ scaled by a factor $1/M$.



a) FREQUENCY SPECTRUM OF A SIGNAL BANDLIMITED TO $\omega = \pi/M$.

b) FREQUENCY SPECTRUM OF DOWNSAMPLED SIGNAL W.R.T ω_x

c) FREQUENCY SPECTRUM OF DOWNSAMPLED SIGNAL W.R.T ω_y.

ALIASING EFFECT.

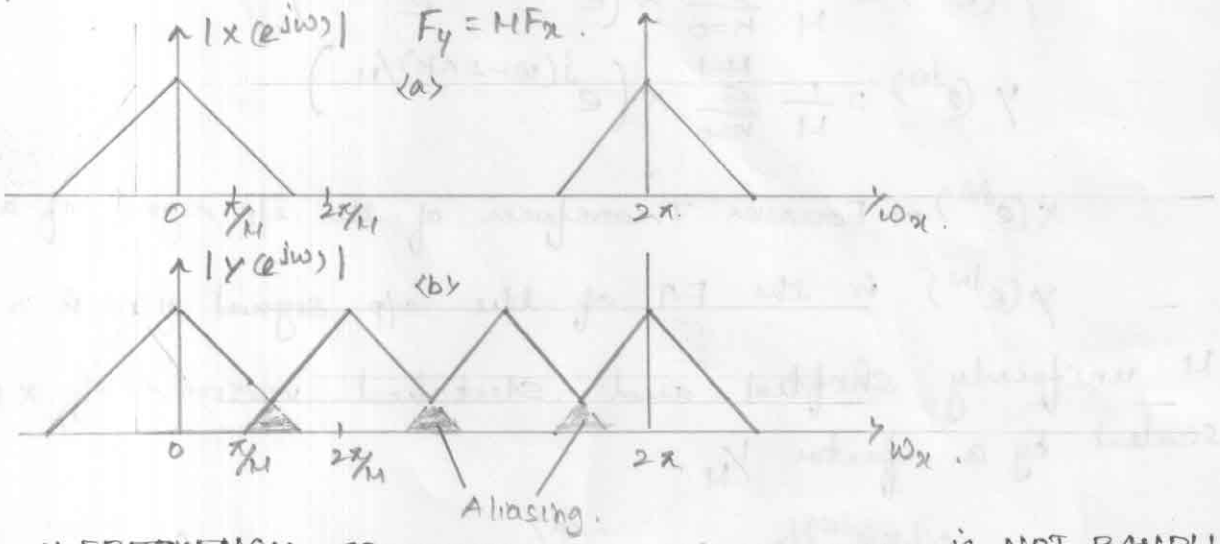
The spectrum obtained after downsampling will overlap if the original spectrum is not bandlimited to $\omega = \pi/M$. This overlap causes aliasing.

www.Vidyarthiplus.com

www.Vidyarthiplus.com

If the signal $x(n)$ is not band limited to $\pm \pi/M$, then a lowpass filter with cut-off frequency π/M is used prior to downsampling. This filter is known as anti-aliasing filter.

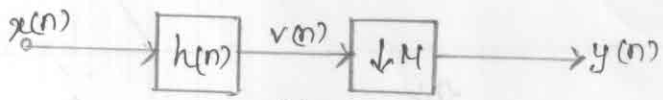
The complet process of filtering and the downsampling is referred to as decimation.



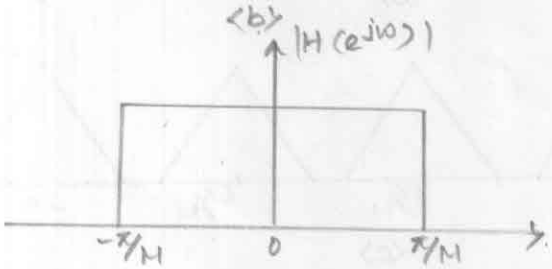
a) FREQUENCY SPECTRUM OF SIGNAL THAT IS NOT BANDLIMITED TO $\omega_c = \pi/M$.

b) FREQUENCY SPECTRUM OF DOWNSAMPLED SIGNAL WITH OVERLAPP

(a)



(b)



a) DECIMATOR WITH AN ANTI-ALIASING FILTER AND A DOWN-SAMPLER

b) IDEAL MAGNITUDE RESPONSE OF ANTI-ALIASING FILTER.

The signal obtained after filtering is

$$v(n) = \sum_{k=-\infty}^{\infty} h(k)\, x(n-k) \quad \text{and}$$

$$y(n) = v(nM) = \sum_{k=-\infty}^{\infty} h(k)\, x(nM-k)$$

www.Vidyarthiplus.com

www.Vidyarthiplus.com

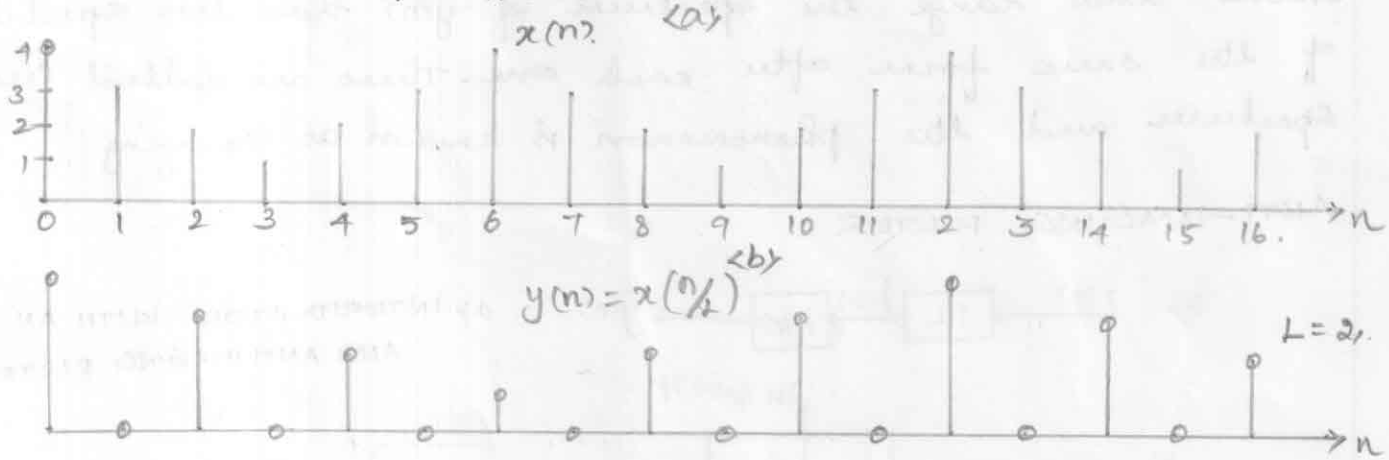
UPSAMPLING

The sampling rate of a discrete-time signal can be increased by a factor L by placing $L-1$ equally spaced zeros b/w each pairs of samples.

Mathematically upsampling is represented by

$$y(n) = \begin{cases} x(n/L) & n = 0, \pm L, \pm 2L, \ldots \\ 0 & \text{otherwise.} \end{cases}$$

UPSAMPLER



SPECTRUM OF THE UPSAMPLED SIGNAL

The z-transform of the signal $y(n)$ is given by

$$Y(z) = \sum_{n=-\infty}^{\infty} y(n) z^{-n}$$

$$= \sum_{n=-\infty}^{\infty} x(n/L) z^{-n}$$

$$= \sum_{n=-\infty}^{\infty} x(m) z^{-Lm} \qquad \because m = n/L \ \& \ n = Lm.$$

$$= \sum_{n=-\infty}^{\infty} x(m) (z^{L})^{-m}$$

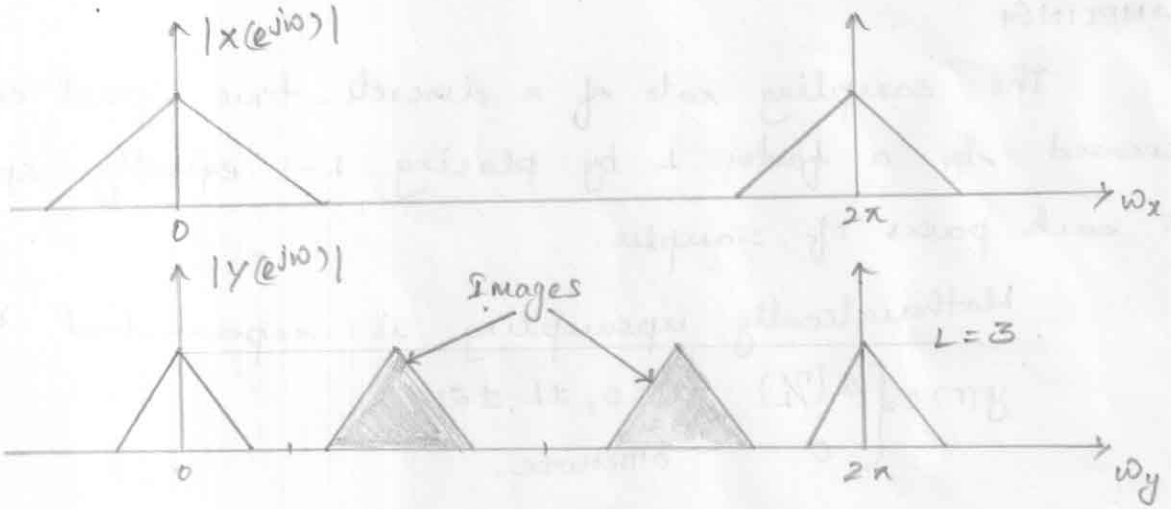
$$Y(z) = X(z^{L})$$

Substituting $z = e^{j\omega}$ in the above equation

$$Y(e^{j\omega}) = X(e^{j\omega L})$$

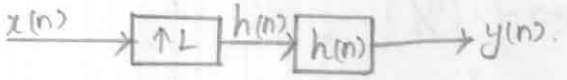
www.Vidyarthiplus.com

www.Vidyarthiplus.com

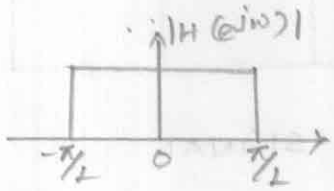


From the above fig. the interval $-\pi/3 < \omega < \pi/3$ the signal $y(n)$ has the same spectrum as $x(n)$ apart from scaling factor. Above this range the spectrum of $y(n)$ has two spectrum of the same form after each one. These are called image spectrum and the phenomenon is known as imaging.

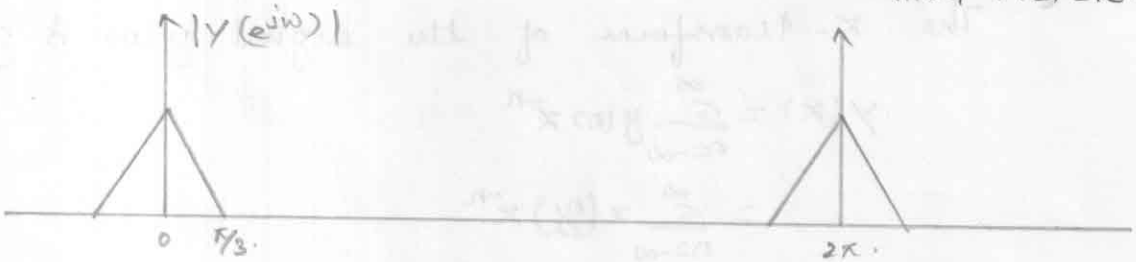
ANTI-IMAGING FILTER

a) 

a) INTERPOLATOR WITH AN UPSAMPLER AND ANTI-IMAGING FILTER

b) 

b) MAGNITUDE RESPONSE OF ANTI-IMAGING FILTER.



SPECTRUM OF THE SIGNAL AT THE O/P OF LPF.

The LPF with a cut-off frequency $\omega_c = \pi/L$ can be used after upsampler. This filter which is used to remove the image spectrum is known as anti-imaging filter.

The Complete process of upsampling and filtering is known as interpolation.

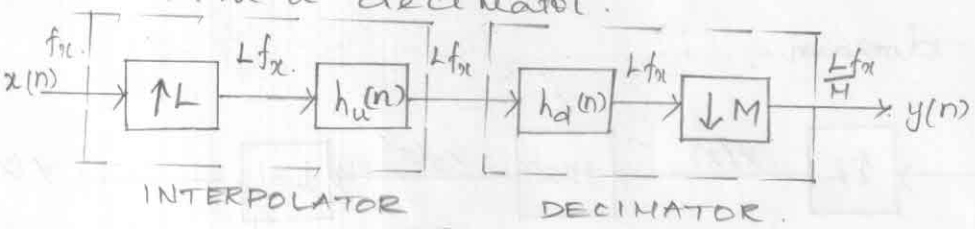
www.Vidyarthiplus.com

www.Vidyarthiplus.com

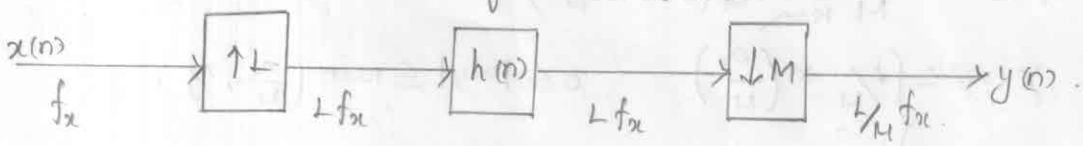
SAMPLING RATE CONVERSION BY A RATIONAL FACTOR (L/M)

Basically we can achieve this sampling rate conversion by first performing interpolation by a factor L and then decimating the o/p of interpolator by a factor M.

In otherwords the sampling rate conversion by rational factor L/M is accomplished by cascading an interpolator with a decimator.



Since $h_u(n)$ and $h_d(n)$ have same sampling rate, we combine them to form $h(n)$.



The importance of performing the interpolation and decimation is to preserve the desired spectral characteristics of $h(n)$.

Here the filters are operated at the same rate namely Lf_x, because of this 2 filters can be combined into a single LPF with impulse response $h(n)$.

The frequency response of $h(n)$ must incorporate the filtering operations for both interpolation and decimation.

The cut-off frequency of LPF is

$$\omega_c = \min\left(\pi/L,\ \pi/M\right)$$

Frequency response, $H(\omega) = \begin{cases} L & 0 \le |\omega| \le \min(\pi/L, \pi/M) \\ 0 & \text{otherwise} \end{cases}$

Here $\omega = \dfrac{2\pi f}{L f_x} = \dfrac{\omega_x}{L}$ $\qquad \because \omega = 2\pi f\ ;\ \omega_s = \dfrac{2\pi f}{f_s}$.

www.Vidyarthiplus.com

www.Vidyarthiplus.com

We have $y(n) = \sum_{k=-\infty}^{\infty} x(k)\, h(nM - kL)$

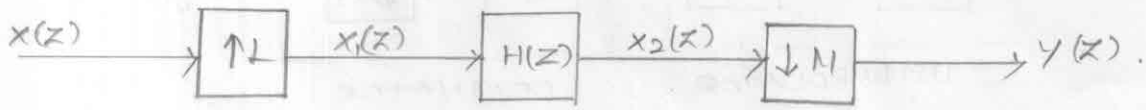
Let $m = nM - kL \Rightarrow k = \frac{nM}{L} - \frac{m}{L} \approx \left\lfloor \frac{nM}{L} \right\rfloor - m$.

$$y(n) = \sum_{m=-\infty}^{\infty} x\left(\left\lfloor \frac{nM}{L} \right\rfloor - m\right) . h\left(nM - \left\lfloor \frac{nM}{L} \right\rfloor L + mL\right)$$

Let $nM - \left\lfloor \frac{nM}{L} \right\rfloor L = nM$ modulo L. $= (nM)_L$.

$$\therefore y(n) = \sum_{m=-\infty}^{\infty} x\left(\left\lfloor \frac{nM}{L} \right\rfloor - m\right) h\left(mL + (nM)_L\right)$$

In Frequency domain.



$X_1(z) = X(z^L) \Rightarrow X_2(z) = H(z)\, X_1(z) = H(z)\, X(z^L) = L \cdot X(z^L)$

$$Y(z) = \frac{1}{M} \sum_{k=0}^{M-1} X_2\left(z^{1/M} W_M^k\right)$$

$$Y(\omega) = \begin{cases} \frac{L}{M} X\left(\frac{\omega_y}{M}\right) & 0 \le |\omega_y| \le \min\left(\frac{\pi}{M}, \frac{\pi}{L}\right) \\ 0 & \text{Otherwise} \end{cases}$$

where $\omega_y = \frac{L}{M} \omega_x$.

www.Vidyarthiplus.com

www.Vidyarthiplus.com

IMAGE ENHANCEMENT

The aim of image enhancement is to improve the interoperability or perception of information in images for human viewers. or to provide better input for other automated image processing techniques.

The image enhancement approaches can be divided into two broad catagories they are

1. Spatial domain approach
2. Frequency domain approach.

SPATIAL DOMAIN APPROACH.

The spatial domain refers to the image plane itself, and approaches in this category are based on direct manipulation of pixels in an image.

Image processing functions in the spatial domain may be expressed as

$$g(x,y) = T[f(x,y)]$$

Where $f(x,y)$ is the input image, $g(x,y)$ is the processed enhanced image and T represents an operator on 'f' defined over some neighborhood of (x,y).

Enhancement at any point in an image depends only on the gray level at that point. This kind of technique is called as "point processing".

www.Vidyarthiplus.com

www.Vidyarthiplus.com

The larger neighbourhood allow a variety of processing functions that go beyond just image enhancement. The neighbourhood of pixels considered about the center pixel (x, y) is called as mask, template or window.

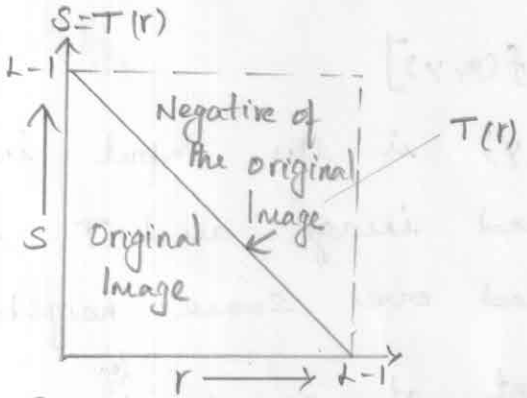
A mask is a small two dimentional array having the size of (3x3), (5x5) and (7x7). The enhancement technique based on this type is called as "Mask processing or filtering".

NEGATIVE OF AN IMAGE.

This image enhancement technique is based only on the intensity of single pixel.

The negative of the digital image is obtained by using the transformation function $s = T[r]$.

Where, 'r' denotes the intensity of pixel in the given original image and 's' denotes the intensity of the pixels in the enhanced image.



GRAY LEVEL TRANSFORMATION FUNCTION

Where L is the number of gray levels.

The idea behind is to reverse the order from black to white so that the intensity of the o/p image decreases as the intensity of the i/p increases.

www.Vidyarthiplus.com

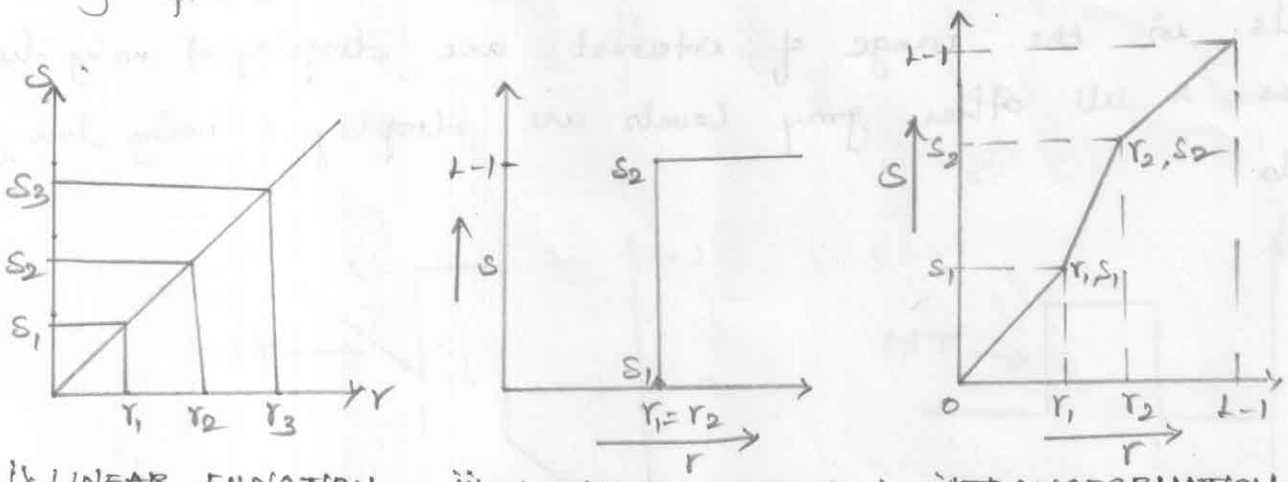
www.Vidyarthiplus.com

CONTRAST STRETCHING

During image acquisition low contrast images may result due to one of the following reasons.

1. Poor illumination
2. Lack of dynamic range in the image sensor
3. Wrong setting of the image sensor.

The idea behind the contrast stretching is to increase the dynamic range of gray levels in the image being processed.



i) LINEAR FUNCTION ii) TRANSFER FUNCTION iii) TRANSFORMATION FUNCTION

CASE (i)

From the fig (i) $r_1 = s_1$, $r_2 = s_2$, $r_3 = s_3$ linearly proportional so both are same. s - processed image, r - original image

CASE (ii)

From the fig (ii) the location of the points (r_1, s_1) and (r_2, s_2) control the shape of the transformation function.

If $r_1 = r_2$, $s_1 = 0$ & $s_2 = L-1$, the transformation becomes the thresholding function and creates a binary image. This kind of transformation is called thresholding function.

CASE (iii)

Intermediate values of r_1, s_1 & r_2, s_2 produces various degree of spread in the gray levels at the o/p image.

contrast stretching is obtained by setting

$(r_1, s_1) = (r_{min}, 0)$, $(r_2, s_2) = (r_{max}, L-1)$

r_{max} - max. gray level, r_{min} - min. gray level.

www.Vidyarthiplus.com

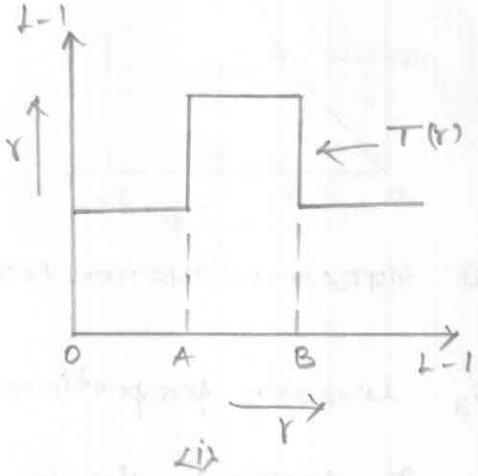
www.Vidyarthiplus.com

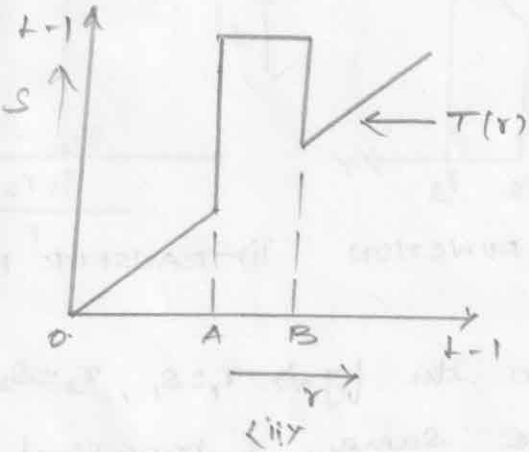
GRAY LEVEL SLICING

There are numerous applications in which highlighting a specific range of gray levels in an image often required.

For example, enhancing the flaws in X-ray images & enhancing features such as masses of water in satellite imagery.

An enhancement technique in which all the gray levels in the range of interest are displayed using high values & all other gray levels are displayed using low gray levels.



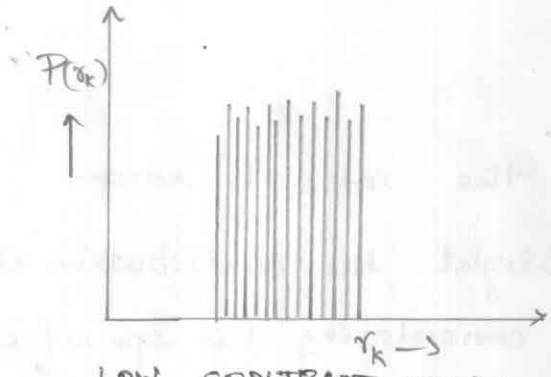


In the first approach, all the gray levels in the range of interest are displayed using low values. The corresponding transformation function used is shown in fig(i) and this results in a binary image.

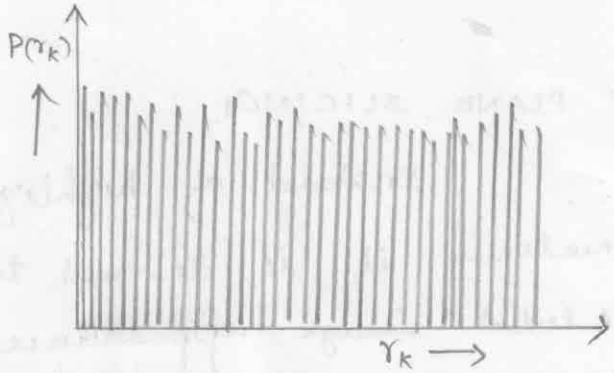
The second approach is based on transformation function shown in fig.(ii) This transfer function brightens the desired range of gray levels but preserve the background and the gray level tonalities in the image.

www.Vidyarthiplus.com

www.Vidyarthiplus.com



LOW CONTRAST IMAGE



HIGH CONTRAST IMAGE

HISTOGRAM EQUALISATION

Equalisation is a process that attempts to spread out the gray levels in an image so that they are evenly distributed across their range.

Histogram equalisation reassigns the brightness value of pixels based on the image histogram.

SPATIAL FILTERING [LOCAL/NEIGHBOURHOOD OPERATION]

Spatial filtering involves passing a weighted mask, or kernal over the image and replacing the original image pixel value corresponding to the centre of the kernal with the sum of the original pixel values in the region corresponding to the kernal multiplied by the kernal weight.

MEAN FILTER OR AVERAGING FILTER OR LOW PASS FILTER

The mean filter replaces each pixel by the average of ~~the~~ all the values in the local neighbourhood. The size of the neighbourhood controls the amount of filtering.

In a spatial averaging operation, each pixel is replaced by a weighted average of its neighbourhood pixels.

The LPF preserves the smooth region in the image and it removes the sharp variations leading to blurring effect.

3 by 3 Low-pass spatial mask $= \frac{1}{9} \times \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$

www.Vidyarthiplus.com

www.Vidyarthiplus.com

BIT PLANE SLICING

Instead of highlighting the intensity ranges sometimes it is desired to highlight the contribution made to total image appearance by considering the specific bits of the image.

Each pixel in the image is represented by 8 bits. Imagine that the image is composed of eight 1 bit planes corresponding to the 8 bits ranging from plane 0 for least significant and plane 7 for the most significant bit.

Plane 0 contains all the lower order bits in the bytes comprising the pixels in the image and Plane 7 contains all the higher order bits.

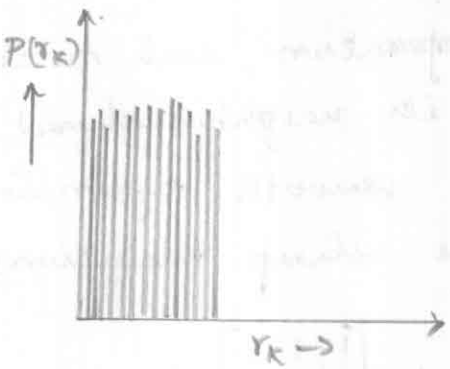
HISTOGRAM

Histogram of a digital image is the probability of occurrence associated with the gray levels in the range 0 to 255.

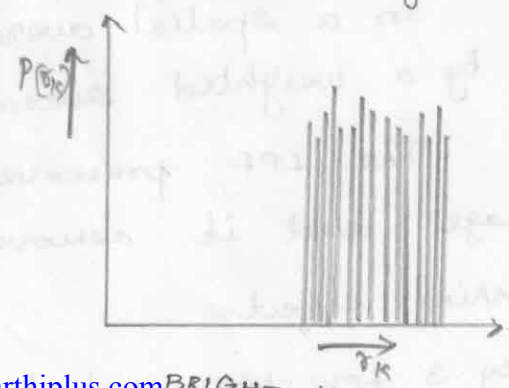
Histogram can be expressed using discrete function

$$P(r_k) = \frac{n_k}{n}$$

Where r_k is the kth gray level, n_k is the number of pixels in the image with that gray level, n is the total number of pixels in the image and $k = 0, 1, 2, \ldots 255$



DARK IMAGE



BRIGHT IMAGE

www.Vidyarthiplus.com

www.Vidyarthiplus.com

IMAGE ENHANCEMENT IN THE FREQUENCY DOMAIN.

Frequency refers to the rate of repetion of some periodic event.

In image processing, spatial frequency refers to the variation of image brightness with its position in space.

Convolution in spatial domain = Multiplication in Frequency domain

$$f(m,n) * h(m,n) = F(k,l) \times H(k,l).$$

BUTTERWORTH LOW PASS FILTER

The transfer function of a two dimensional Butterworth low-pass filter is given by

$$H(k,l) = \frac{1}{1+\left[\frac{\sqrt{k^2+l^2}}{D_0}\right]^2}$$

Where, $n \rightarrow$ Order of the filter

$D_0 \rightarrow$ cut-off frequency.

BUTTERWORTH HIGHPASS FILTER

The transfer function of a two-dimensional Butterworth filter is given by

$$H(k,l) = \frac{1}{1+\left(\frac{D_0}{\sqrt{k^2+l^2}}\right)^{2n}}$$

Where, $n \rightarrow$ Order of the filter

$D_0 \rightarrow$ cut-off frequency.

www.Vidyarthiplus.com

www.Vidyarthiplus.com

WEIGHTED AVERAGE FILTER

The mask of a weighted average filter is given by

$$\begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}.$$

(Here average is taken from sum of all the neighbouring pixel values in the mask)

GAUSSIAN FILTER

Gaussian filters are a class of linear smoothing filters with the weights chosen according to the shape of a Gaussian function.

The Gaussian filter in the continuous space is given by,

$$h(m,n) = \left(\frac{1}{\sqrt{2\pi}\sigma} e^{-m^2/2\sigma^2}\right) \times \left(\frac{1}{\sqrt{2\pi}\sigma} e^{-n^2/2\sigma^2}\right)$$

The above expression shows that Gaussian filter is separable.

MEDIAN FILTER

A Median filter smoothens the image by utilising the median of the neighbourhood.

SPATIAL DOMAIN HIGH-PASS FILTERING OR IMAGE SHARPENING

The main aim of image sharpening is to highlight fine details in the image.

Image sharpening is used to enhance the high frequency components.

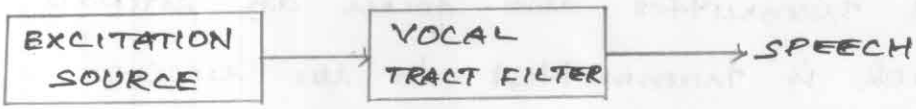
The spatial mask which performs image sharpening is given as

$$\frac{1}{9} \times \begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix}$$

www.Vidyarthiplus.com

www.Vidyarthiplus.com

SPEECH COMPRESSION



A MODEL FOR SPEECH SYNTHESIS

Here the excitation source corresponds to the sound generation, and the vocal tract filter models the vocal tract.

In order to generate a specific fragment of speech, we have to generate a sequence of appropriate vocal tract approximations.

At the transmitter, the speech is divided into segments. Each segment is analyzed to determine an excitation signal and the parameters of the vocal tract filter.

In some of the schemes, a model for the excitation signal is transmitted to the receiver. The excitation signal is then synthesized at the receiver and used to drive the vocal tract filter.

In other schemes, the excitation signal itself is obtained using an analysis-by-synthesis approach. This signal is then used by the vocal tract filter to generate the speech signal.

THE CHANNEL VOCODER

In the channel vocoder, each segment of input speech is analyzed using a bank of band-pass filters called the analysis filters.

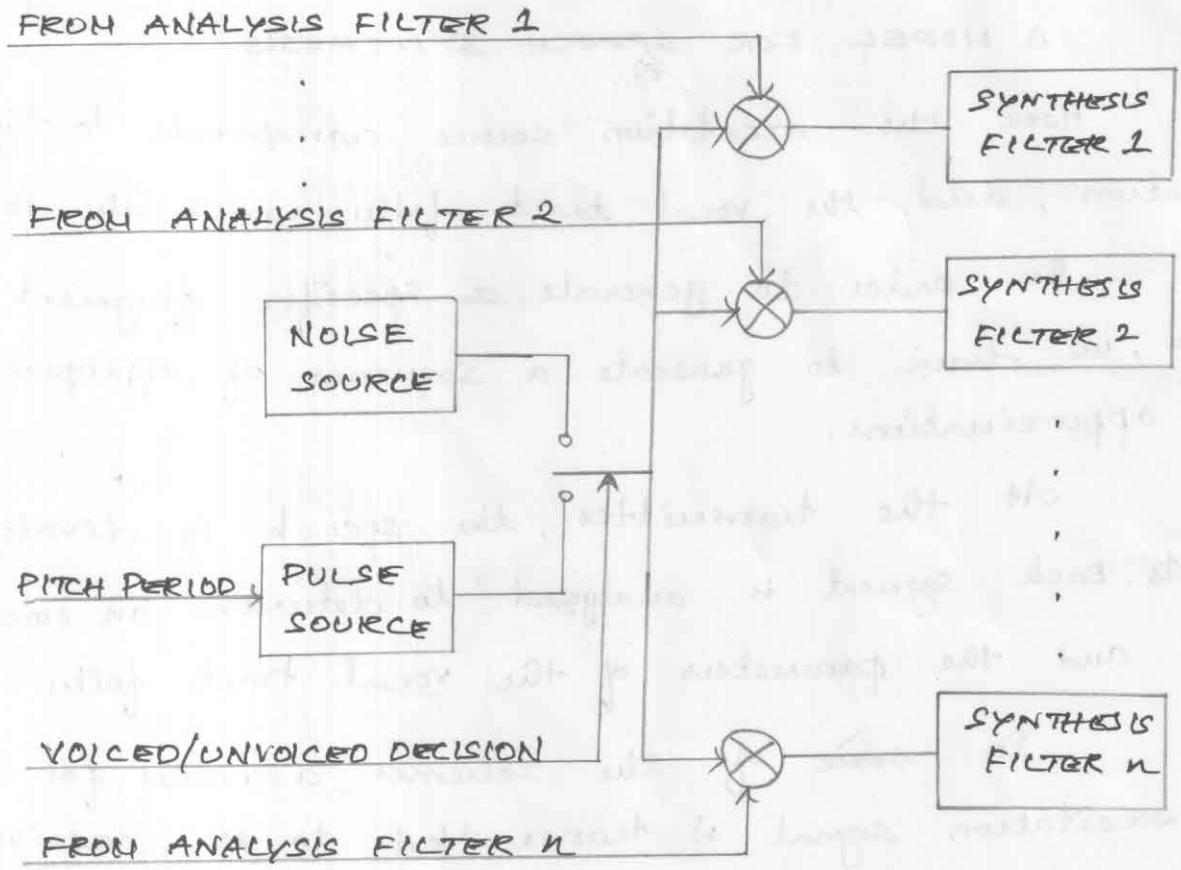
The energy at the o/p of each filter is estimated at fixed intervals and transmitted to the receiver.

The period of the fundamental harmonic is called the pitch period.

www.Vidyarthiplus.com

www.Vidyarthiplus.com

The transmitter also forms an estimate of the pitch period, which is transmitted to the receiver.



THE CHANNEL VOCODER RECEIVER

Voiced sounds tend to have a pseudoperiodic structure. ex: sound /a/ /e/ /o/.

Unvoiced sounds tend to have a noselike structure, which is the /s/ sound in the word test.

At the receiver, the vocal tract filter is implemented by a bank of band pass filter. The bank of filter at the receiver, known as the synthesis filters, is identical to the bank of analysis filters.

Based on whether the speech segment was deemed to be voiced or unvoiced, either a pseudonoise source or a periodic pulse generator is used as the input to the synthesis filter bank.

www.Vidyarthiplus.com

www.Vidyarthiplus.com

The period of the pulse i/p is determined by the pitch estimate obtained for the segment being synthesized at the transmitter.

The channel vocoder matches the frequency profile of the i/p speech. There is no attempt to reproduce the speech samples per sec.

However, not all frequency components of speech are equally important. In fact, as the vocal tract is a tube of non uniform cross section, it resonates at a number of different frequencies. These frequencies are known as formants. The formant values change with different sounds.

Eg: Formant range : 200-800 Hz for Male speaker and 250-1000 Hz for a Female speaker.

Formant vocoders, which transmit an estimate of the formant values and an estimate of the bandwidth of each formant.

In the voice-excited channel vocoder, the voice is first filtered using a narrow-band low pass filter.

The o/p of the LPF is sampled and transmitted to the receiver.

At the receiver, this low pass filtered signal is passed through a non linearity to generate higher order harmonics that together with the LP signal, are used as the excitation signal.

Voice excitation removes the problem of pitch extraction. It also removes the necessity for declaring every segment either voiced/unvoiced.

Unfortunately, the increase in quality is reflected in the high cost of transmitting the low pass filtered speech signal.

www.Vidyarthiplus.com

5.7 Musical Sound Processing

The musical programs are produced in two stages. In the first stage the sound from each individual singer or instrument is recorded in acoustically inert studio on single track of multitrack tape. In the second stage, the special audio effects are added to these individual sounds. Then these individual sounds are combined in the mix-down system to generate stereo recording. The special audio effects can be generated by DSP as discussed next.

Generation of echo effects

The echoes are generated by delay units. For example direct sound and its single echo can be generated with the help of FIR filter having following difference equation.

$$y(n) = x(n) + \alpha\, x(n-R) \qquad \text{... (5.7.1)}$$

Hence $x(n)$ is the direct sound and $x(n-R)$ is the echo delayed by 'R' sampling periods. The amplitude of echo is 'α'. The transfer function of this filter will be,

$$H(z) = 1 + \alpha z^{-R} \qquad \text{... (5.7.2)}$$

Fig. 5.7.1 shows the realization of such filter.

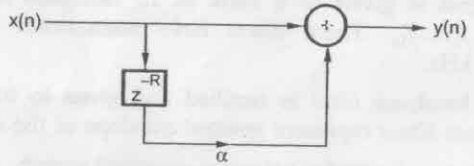


Fig. 5.7.1 A single echo filter

The multiple echoes which are spaced 'R' sampling periods apart and exponentially decaying amplitudes can be generated with the help of FIR filter having transfer function as follows.

$$H(z) = 1 + \alpha z^{-R} + \alpha^2 z^{-2R} + \ldots\ldots + \alpha^{N-1} z^{-(N-1)R}$$

$$= \frac{1-\alpha^N z^{-NR}}{1-\alpha z^{-R}} \qquad \text{... (5.7.3)}$$

Thus the above filter can be realized by IIR filter also. An infinite number of echoes having exponentially decaying amplitudes and spaced at 'R' sampling periods can be generated by IIR filter having following transfer function.

$$H(z) = \frac{z^{-R}}{1-\alpha z^{-R}}, \quad |\alpha| < 1 \qquad \text{... (5.7.4)}$$

www.Vidyarthiplus.com

Fig. 5.7.2 shows realization of this filter.

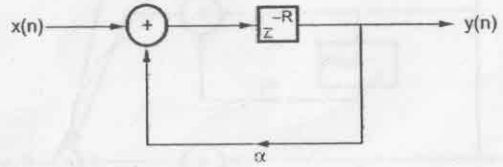


Fig. 5.7.2 IIR filter for generating infinite number of echoes

Generation of reverberation

The echoes generated by schemes discussed just now do not provide natural sounding. This is because their magnitude is not constant at all frequencies and the output echo density is much lower than that observed in a real room. This causes a fluttering of the composite sound.

Hence reverberation is used that generates flutterfree sounds. The reverberation needs 1000 echoes per second. The allpass filter of Fig. 5.7.3 can be used to generate reverberation.

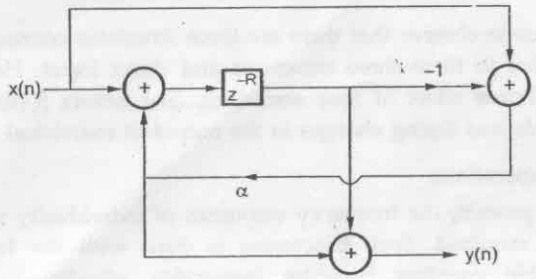


Fig. 5.7.3 Structure of an allpass reverberator

The transfer function of the allpass reverberator shown in above figure is given as,

$$H(z) = \frac{\alpha + z^{-R}}{1+\alpha z^{-R}}, \quad |\alpha| < 1 \qquad \text{... (5.7.5)}$$

Since this is an allpass filter, the magnitude of all the components remain same. Hence it provides more natural sounding.

Generation of chorus effect

The chorus effect is obtained when several musicians play the same musical piece at the same time. The chorus effect is generated because of amplitude and timing differences of individual musicians. Such effect can be generated with the help of chorus generator. Fig. 5.7.4 shows the structure for generating chorus effect.

www.Vidyarthiplus.com

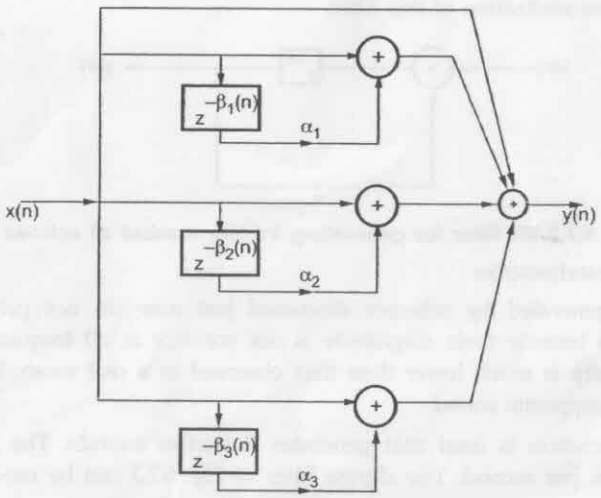


Fig. 5.7.4 Generation of chorus effect

In the above structure observe that there are three structures connected in parallel. The output $y(n)$ is due to these three structures and direct input. Hence the above structure generates chorus effect of four musicians. The factors $\beta_i(n)$ are randomly varied to get amplitude and timing changes in the output of individual structures.

Frequency domain operations

In the mix-down process, the frequency responses of individually recorded pieces of sound are to be modified. Such processing is done with the help of graphic equalizer. The graphic equaliser provides bass-treble adjustments. The graphic equalizer can be implemented with the help of cascaded digital filters.

5.8 Speech Compression

The compression of speech can be achieved by extracting significant features. These significant features are then used to synthesize or decompress the speech. Hence speech analysis by a vocoder becomes the compression and synthesis by a vocoder becomes decompression.

· Vocoder extracts the spectral envelope of speech and information regarding voicing and pitch. This data is coded and transmitted. The synthesizer generates speech from the received data.

www.Vidyarthiplus.com

5.8.1 Vocoder - Analyzer (Speech Compression)

Fig. 5.8.1 shows the block diagram of analyzer of a vocoder.

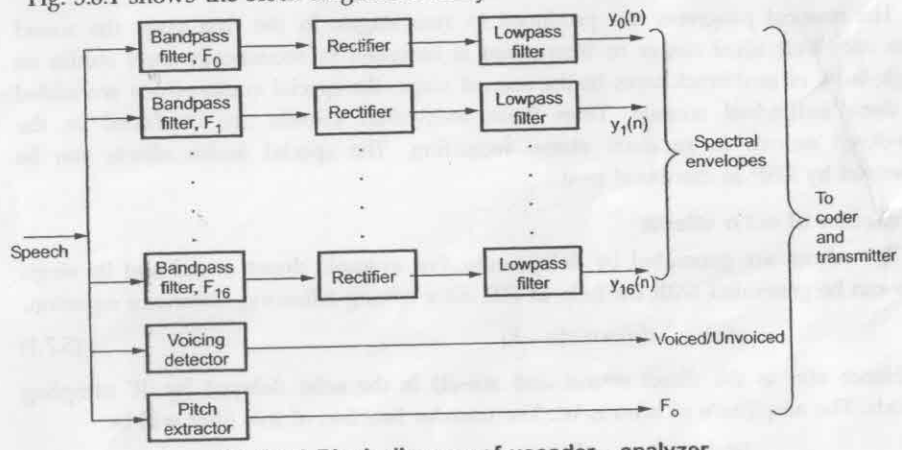


Fig. 5.8.1 Block diagram of vocoder - analyzer

Operation

- The speech signal is given to a bank of 16 bandpass filters centered at frequencies $f_0, f_1, \ldots f_{16}$. These filters have nonuniform frequency bands varying upto 3 kHz.
- The output of bandpass filter is rectified and given to lowpass filter. The output of lowpass filters represent spectral envelope of the signal.
- The voicing detector checks for voiced or unvoiced speech.
- The pitch extractor determine fundamental frequency F_o of the speech signal.

5.8.2 Vocoder - Synthesizer (Speech Decompression)

Fig. 5.8.2 shows the block diagram of the synthesizer of vocoder.

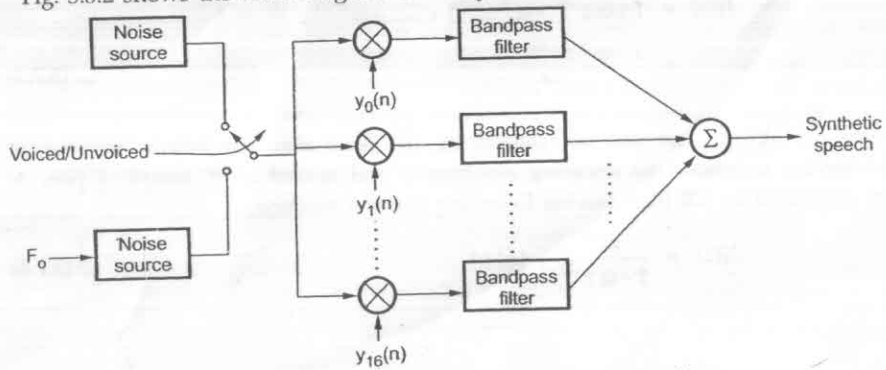


Fig. 5.8.2 Block diagram of vocoder - synthesizer

www.Vidyarthiplus.com

Operation

- The excitation can be created from noise source or pulse source. The voiced / unvoiced signal received from the transmitter is used to select the excitation source.
- The fundamental frequency of the pulse source is controlled by F_0, i.e. information of pitch received from transmitter.
- The lowpass spectrum envelop signals $y_o(n), y_1(n), \ldots y_{16}(n)$ are used to modulate the outputs of bandpass filters. These envelope signals are received from transmitter.
- The synthetic speech is obtained by adding modulated outputs of all bandpass filters.

Advantages

i) Vocoders provide transmission of speech at reduced bit rate with reasonably good quality.

ii) The bit rate 2400 to 96 bits/sec is adequate for vocoders.

5.8.3 Adaptive Filters

Principle : The coefficients of the filter are changed automatically according to the changes in input signal. This means the filtering characteristics of the adaptive filter are changed or adapted according to the changes in input signal.

Block diagram

Fig. 5.8.3 shows the block diagram of adaptive filter which is used as simple noise canceller. The signal to be filtered is noisy, i.e. $y(n)$. It consists of signal $s(n)$ plus

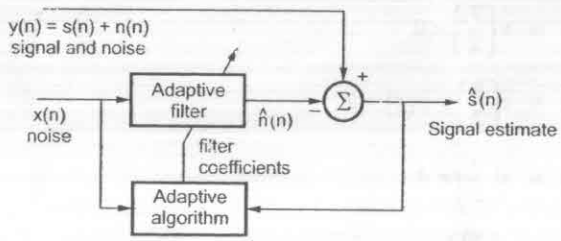


Fig. 5.8.3 Adaptive filter

noise $n(n)$. Only noise from the same signal source is $x(n)$ This noise is given to adaptive filter. The characteristics of the adaptive filter are changed by adaptive algorithm. The adaptive algorithm modifies the filter coefficients depending upon input noise and output signal estimate. The adaptive filter produces an estimate of noise $\hat{n}(n)$. This estimate is subtracted from noisy signal i.e.,

www.Vidyarthiplus.com

$$\hat{s}(n) = y(n) - \hat{n}(n)$$
$$= s(n) + n(n) - \hat{n}(n)$$

Thus output signal $\hat{s}(n)$ is the estimate of noise free signal. The adaptive filter and adaptive algorithm combinely forms an adaptive filter.

Advantages

i) Adaptive filters extract the signal from noisy signals whose characteristics vary with time.

ii) Adaptive filters extract the signals where signal and noise have the same spectrum.

iii) The filtered signal contains minimum distortion.

iv) Adaptive filters have built in flexibility.

Applications

i) Filtering of artefacts from the human EEG.

ii) Filtering of jamming signal in spread spectrum communication.

iii) Adaptive echo cancellation on telephone channels.

iv) Radar signal processing

v) Navigational systems.

vi) Adaptive equalization of communication channels.

Adaptive filtering algorithms

i) Least mean square algorithm

ii) Recursive least squares algorithm.

These algorithms are used to adaptively calculate filter coefficients.

5.8.4 Image Enhancement

Principle : Certain features of the image are emphasized for subsequent analysis and display.

Image enhancement techniques : Following is the list of important image enhancement methods :

i) Contrast and edge enhancement

ii) Pseudo colouring

iii) Noise filtering

iv) Sharpening

v) Magnifying

www.Vidyarthiplus.com

vi) Histogram sliding and stretching

vii) Brightness slicing

viii) Photometric correction

DSP techniques used for image enhancement

Following DSP tools are used for image enhancement :

i) Local neighbourhood operations as in convolution

ii) Transform operations as in DFT

iii) Mapping operations as in pseudo colouring and gray level mapping.

Applications

i) Feature extraction in an image

ii) Image analysis

iii) Visual information display.

5.9 Short Answered Questions [2 Marks Each]

Q.1 *Represent decimal number 0.69 in fixed point representation of length N=6.* **[Nov./Dec.-2010, IT]**

Ans. : $(0.69)_{10} = (0.1011000010...)_2$

For $N = 6$ above number will be,

$$(0.69)_{10} = (\underset{\uparrow}{0}.10110)_2$$

Indicates positive number

This is signed magnitude format.

Q.2 *Prove that up sampling by a factor M is time varying system.* **[Nov./Dec.-2010, IT]**

Ans. : Up sampling is given by

$$y(m) = \sum_{k=-\infty}^{\infty} h(m-kM)\, x(k)$$

Here note that 'm' and 'k' are different indices. The impulse response $h(m-kM)$ is the function of 'm' and 'k'. Thus up sampling operation is time varying system.

Q.3 *State a few applications of adaptive filter.* **[Nov./Dec.-2010, IT]**

Ans. :

1) Filtering of jamming signal in spread spectrum communication.

2) Adaptive echo cancellation on telephone channels.

3) Adaptive equalization of communication channels.

4) Radar signal processing.

University Questions with Answers

Q.1 *A signal x(n) = {6, 1, 5, 7, 2, 1}*
Find :
1) x(n/2) 2) x(2n). **Nov./Dec.-2010, 8 Marks, IT**

Ans. : 1) $y(n) = x\left(\frac{n}{2}\right)$

Here

$$y(0) = x(0) = 6$$

$$y(1) = x\left(\frac{1}{2}\right) = 0$$

$$y(2) = x\left(\frac{2}{2}\right) = x(1) = 1$$

$$y(3) = x\left(\frac{3}{2}\right) = 0$$

$$y(4) = x\left(\frac{4}{2}\right) = x(2) = 5$$

$$y(5) = x\left(\frac{5}{2}\right) = 0$$

$$y(6) = x\left(\frac{6}{2}\right) = x(3) = 7$$

$$y(7) = x\left(\frac{7}{2}\right) = 0$$

$$y(8) = x\left(\frac{8}{2}\right) = x(4) = 2$$

$$y(9) = x\left(\frac{9}{2}\right) = 0$$

$$y(10) = x\left(\frac{10}{2}\right) = x(5) = 1$$

Thus, $y(n) = x\left(\frac{n}{2}\right) = \{6, 0, 1, 0, 5, 0, 7, 0, 2, 0, 1\}$

www.Vidyarthiplus.com
www.Vidyarthiplus.com

2) $y(n) = x(2n)$

$$y(0) = x(0) = 6$$
$$y(1) = x(2\times1) = x(2) = 5$$
$$y(2) = x(2\times2) = x(4) = 2$$
$$y(3) = x(2\times3) = x(6) = 0$$

Thus, $y(n) = x(2n) = \{6, 5, 2, 0\}$

Q.2 *Explain any one application using multirate processing of signals.*

Nov./Dec.-2010, 8 Marks, IT

Ans. : Necessity of subband coding : Speech energy is not uniformly distributed over the complete frequency range. More energy is contained in lower frequency bands and less energy is contained in higher frequency bands. Hence lower frequency bands are coded with more number of bits and high frequency bands are coded with less number of bits.

Definition of subband coding : The speech signal is divided into several frequency bands. These frequency band do not overlap each other. They are called subbands. Each subband is encoded separately.

Block diagram of subband coding : Fig. 1 shows the block diagram of subband coding

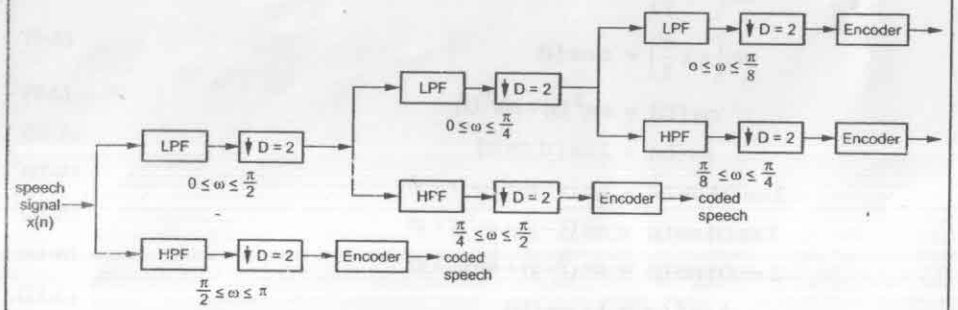


Fig. 1 Block diagram of subband coding

- The speech signal containing the frequencies from 0 to π radians is passed through a pair of lowpass filter and high pass filter.

- These lowpass and high pass filters split the signal into two non overlapping frequency bands. These bands are $0 \le \omega \le \frac{\pi}{2}$ and $\frac{\pi}{2} \le \omega \le \pi$ Fig. 2 shows these bands.

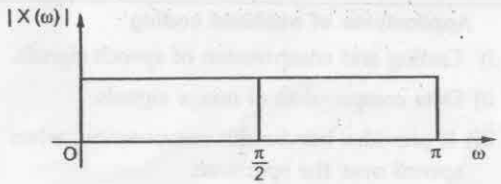


Fig. 2 Frequency bands

- The lower frequency band is further split into two frequency bands i.e. $0 \le \omega \le \frac{\pi}{4}$ and $\frac{\pi}{4} \le \omega \le \frac{\pi}{2}$. This is shown in Fig. 3. Note that higher frequency band of $\frac{\pi}{2} \le \omega \le \pi$ is not split further. It is directly encoded since it contains less energy.

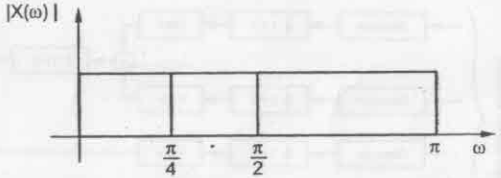


Fig. 3

- The lowest frequecny band is further split into $0 \le \omega \le \frac{\pi}{8}$ and $\frac{\pi}{8} \le \omega \le \frac{\pi}{4}$. Thus the signal is split into four frequency bands as shown in Fig. 4.

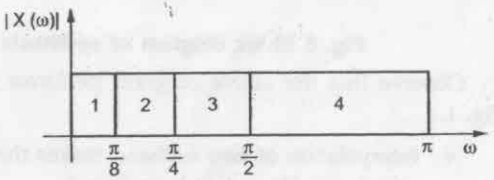


Fig. 4

- The subbands are decimated immediately after they are filtered. This is because, the frequency range is reduced after filtering. Hence the sampling frequency can also be reduced. In Fig. 1 observe that decimation is performated after enery stage of splitting. This requires less number of bits for encoding .

- **Filters used for subband coding :** Fig. 5 shows the characteristic of low pass and high pass filters used for subband coding. It is the characteristic of quadrature mirror filter (QMF). These are physically realizable filters. They eliminate the aliasing resulting due to decimation.

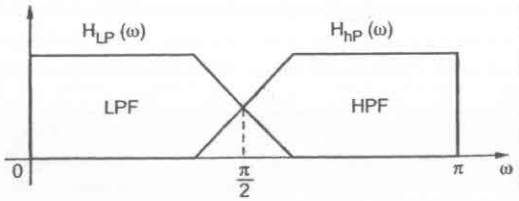


Fig. 5 QMF characteristic

www.Vidyarthiplus.com
www.Vidyarthiplus.com

Applications of subband coding

i) Coding and compression of speech signals.

ii) Data compression of image signals.

iii) It provides bandwidth compression, when the energy of the signal is unevenly spread over the spectrum.

Synthesis of subband coded signal

Fig. 6 shows the block diagram of synthesis of subband coded signal.

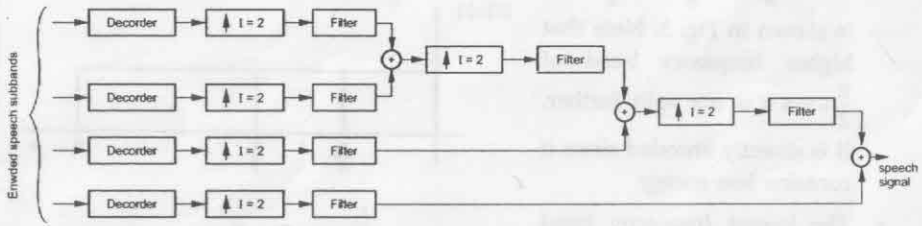


Fig. 6 Block diagram of synthesis of subband coded signal

Observe that the above diagram performs exactly the reverse function of that of Fig. 1.

- Interpolation of two subband makes their sampling rates equal. Then the two signals are filtered and combined.

Q.3 *Write short note on adaptive filter.* **Nov./Dec.-2010, 8 Marks, IT**

Ans. : Refer section 5.8.3.

Q.4 *Write short note image enhancement.* **Nov./Dec.-2010, 8 Marks, IT**

Ans. : Refer section 5.8.4.

□□□

www.Vidyarthiplus.com

A Important Mathematical Relations

Trigonometric Relations :

$$\sin(x) = \frac{e^{jx} - e^{-jx}}{2j} \quad \text{(A-1)}$$

$$\cos(x) = \frac{e^{jx} + e^{-jx}}{2} \quad \text{(A-2)}$$

$$\tan(x) = \frac{\sin(x)}{\cos(x)} = \frac{e^{jx} - e^{-jx}}{j(e^{jx} + e^{-jx})} \quad \text{(A-3)}$$

$$e^{\pm jx} = \cos(x) \pm j\sin(x) \text{ (Euler's theorem)} \quad \text{(A-4)}$$

$$\cos(x \pm y) = \cos(x)\cos(y) \mp \sin(x)\sin(y) \quad \text{(A-5)}$$

$$\sin(x \pm y) = \sin(x)\cos(y) \pm \cos(x)\sin(y) \quad \text{(A-6)}$$

$$\cos\left(x \pm \frac{\pi}{2}\right) = \mp \sin(x) \quad \text{(A-7)}$$

$$\sin\left(x \pm \frac{\pi}{2}\right) = \pm \cos(x) \quad \text{(A-8)}$$

$$\cos(2x) = \cos^2(x) - \sin^2(x) \quad \text{(A-9)}$$

$$\sin(2x) = 2\sin(x)\cos(x) \quad \text{(A-10)}$$

$$2\cos(x)\cos(y) = \cos(x-y) + \cos(x+y) \quad \text{(A-11)}$$

$$2\sin(x)\sin(y) = \cos(x-y) - \cos(x+y) \quad \text{(A-12)}$$

$$2\sin(x)\cos(y) = \sin(x-y) + \sin(x+y) \quad \text{(A-13)}$$

$$2\cos^2(x) = 1 + \cos(2x) \quad \text{(A-14)}$$

$$2\sin^2(x) = 1 - \cos(2x) \quad \text{(A-15)}$$

$$4\cos^3(x) = 3\cos(x) + \cos(3x) \quad \text{(A-16)}$$

$$4\sin^3(x) = 3\sin(x) - \sin(3x) \quad \text{(A-17)}$$

$$8\cos^4(x) = 3 + 4\cos(2x) + \cos(4x) \quad \text{(A-18)}$$

$$8\sin^4(x) = 3 - 4\cos(2x) + \cos(4x) \quad \text{(A-19)}$$

www.Vidyarthiplus.com

14.4 The Basic Subband Coding Algorithm

The basic subband coding system is shown in Figure 14.7.

14.4.1 Analysis

The source output is passed through a bank of filters, called the analysis filter bank, which covers the range of frequencies that make up the source output. The passbands of the filters can be nonoverlapping or overlapping. Nonoverlapping and overlapping filter banks are shown in Figure 14.8. The outputs of the filters are then subsampled.

The justification for the subsampling is the Nyquist rule and its generalization, which tells us that we only need twice as many samples per second as the range of frequencies. This means that we can reduce the number of samples at the output of the filter because the range of frequencies at the output of the filter is less than the range of frequencies at the input to the filter. This process of reducing the number of samples is called *decimation*,[1] or *downsampling*. The amount of decimation depends on the ratio of the bandwidth of the filter output to the filter input. If the bandwidth at the output of the filter is $1/M$ of the bandwidth at the input to the filter, we would decimate the output by a factor of M by keeping every Mth sample. The symbol $M\downarrow$ is used to denote this decimation.

Once the output of the filters has been decimated, the output is encoded using one of several encoding schemes, including ADPCM, PCM, and vector quantization.

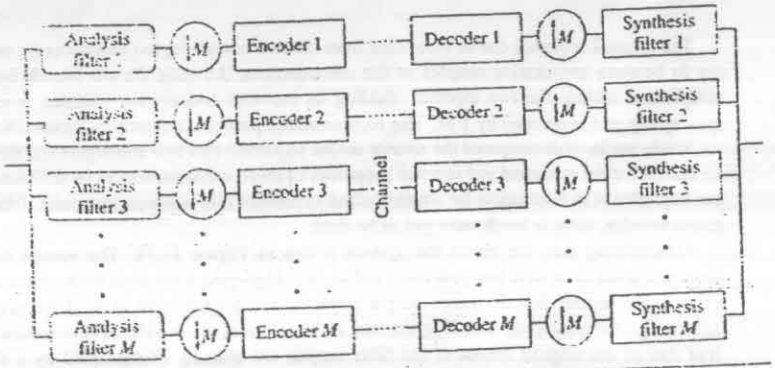


FIGURE 14.7 Block diagram of the subband coding system.

[1] The word *decimation* has a rather bloody origin. During the time of the Roman empire, if a legion broke ranks and ran during battle, its members were lined up and every tenth person was killed. This process was called decimation.

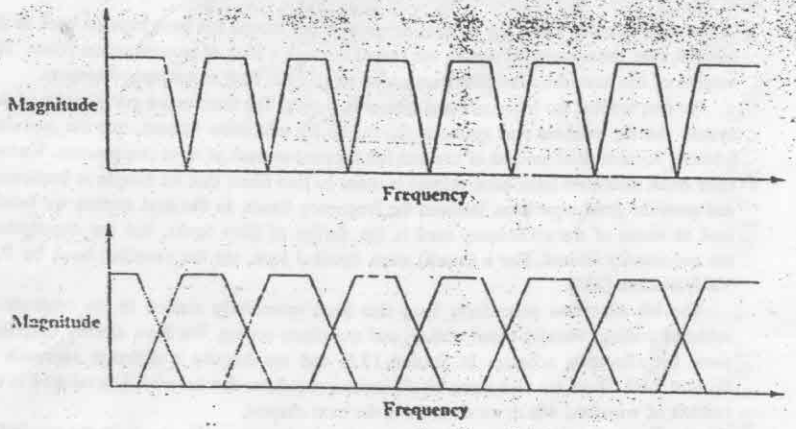


FIGURE 14.8 Nonoverlapping and overlapping filter banks.

14.4.2 Quantization and Coding

Along with the selection of the compression scheme, the allocation of bits between the subbands is an important design parameter. Different subbands contain differing amounts of information. Therefore, we need to allocate the available bits among the subbands according to some measure of the information content. There are a number of different ways we could distribute the available bits. For example, suppose we were decomposing the source output into four bands and we wanted a coding rate of 1 bit per sample. We could accomplish this by using 1 bit per sample for each of the four bands. On the other hand, we could simply discard the output of two of the bands and use 2 bits per sample for the two remaining bands. Or, we could discard the output of three of the four filters and use 4 bits per sample to encode the output of the remaining filter.

This *bit allocation* procedure can have a significant impact on the quality of the final reconstruction, especially when the information content of different bands is very different.

If we use the variance of the output of each filter as a measure of information, and assume that the compression scheme is scalar quantization, we can arrive at several simple bit allocation schemes (see Section 13.5). If we use a slightly more sophisticated model for the outputs of the filters, we can arrive at significantly better bit allocation procedures (see Section 14.9).

14.4.3 Synthesis

The quantized and coded coefficients are used to reconstruct a representation of the original signal at the decoder. First, the encoded samples from each subband are decoded at the receiver. These decoded values are then upsampled by inserting an appropriate number of

www.Vidyarthiplus.com
www.Vidyarthiplus.com

0s between samples. Once the number of samples per second has been brought back to the original rate, the upsampled signals are passed through a bank of reconstruction filters. The outputs of the reconstruction filters are added to give the final reconstructed outputs.

We can see that the basic subband system is simple. The three major components of this system are the *analysis and synthesis filters*, the *bit allocation* scheme, and the *encoding* scheme. A substantial amount of research has focused on each of these components. Various filter bank structures have been studied in order to find filters that are simple to implement and provide good separation between the frequency bands. In the next section we briefly look at some of the techniques used in the design of filter banks, but our descriptions are necessarily limited. For a (much) more detailed look, see the excellent book by P.P. Vaidyanathan [200].

The bit allocation procedures have also been extensively studied in the contexts of subband coding, wavelet-based coding, and transform coding. We have already described some bit allocation schemes in Section 13.5, and we describe a different approach in Section 14.9. There are also some bit allocation procedures that have been developed in the context of wavelets, which we describe in the next chapter.

The separation of the source output according to frequency also opens up the possibility for innovative ways to use compression algorithms. The decomposition of the source output in this manner provides inputs for the compression algorithms, each of which has more clearly defined characteristics than the original source output. We can use these characteristics to select separate compression schemes appropriate to each of the different inputs.

Human perception of audio and video inputs is frequency dependent. We can use this fact to design our compression schemes so that the frequency bands that are most important to perception are reconstructed most accurately. Whatever distortion there has to be is introduced in the frequency bands to which humans are least sensitive. We describe some applications to the coding of speech, audio, and images later in this chapter.

Before we proceed to bit allocation procedures and implementations, we provide a more mathematical analysis of the subband coding system. We also look at some approaches to the design of filter banks for subband coding. The analysis relies heavily on the Z-transform concepts introduced in Chapter 12 and will primarily be of interest to readers with an electrical engineering background. The material is not essential to understanding the rest of the chapter; if you are not interested in these details, you should skip these sections and go directly to Section 14.9.

14.5 Design of Filter Banks ★

In this and the following starred section we will take a closer look at the analysis, downsampling, upsampling, and synthesis operations. Our approach follows that of [201]. We assume familiarity with the Z-transform concepts of Chapter 12. We begin with some notation. Suppose we have a sequence $x_0, x_1, x_2, \ldots$. We can divide this sequence into two subsequences: $x_0, x_2, x_4, \ldots$ and $x_1, x_3, x_5, \ldots$ using the scheme shown in Figure 14.9, where z^{-1} corresponds to a delay of one sample and $\downarrow M$ denotes a subsampling by a factor of M. This subsampling process is called *downsampling* or *decimation*.

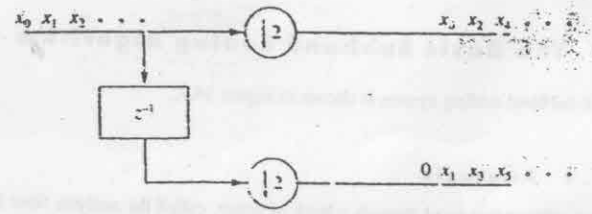


FIGURE 14.9 Decomposition of an input sequence into its odd and even components.

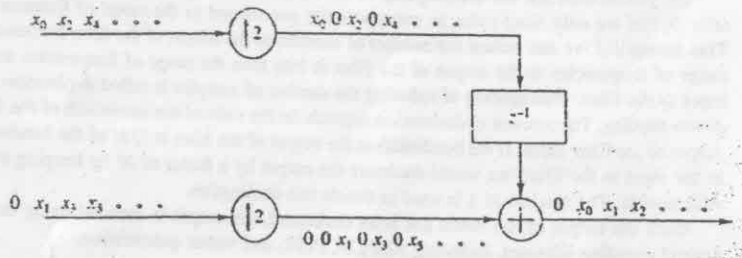


FIGURE 14.10 Reconstructing the input sequence from its odd and even components.

The original sequence can be recovered from the two downsampled sequences by inserting 0s between consecutive samples of the subsequences, delaying the top branch by one sample and adding the two together. Adding 0s between consecutive samples is called *upsampling* and is denoted by $\uparrow M$. The reconstruction process is shown in Figure 14.10.

While we have decomposed the source output sequence into two subsequences, there is no reason for the statistical and spectral properties of these subsequences to be different. As our objective is to decompose the source output sequences into subsequences with differing characteristics, there is much more yet to be done.

Generalizing this, we obtain the system shown in Figure 14.11. The source output sequence is fed to an ideal low-pass filter and an ideal high-pass filter, each with a bandwidth of $\pi/2$. We assume that the source output sequence had a bandwidth of π. If the original source signal was sampled at the Nyquist rate, as the output of the two filters have bandwidths half that of the original sequence, the filter outputs are actually oversampled by a factor of two. We can, therefore, subsample these signals by a factor of two without any loss of information. The two bands now have different characteristics and can be encoded differently. For the moment let's assume that the encoding is performed in a lossless manner so that the reconstructed sequence exactly matches the source output sequence.

www.Vidyarthiplus.com
www.Vidyarthiplus.com

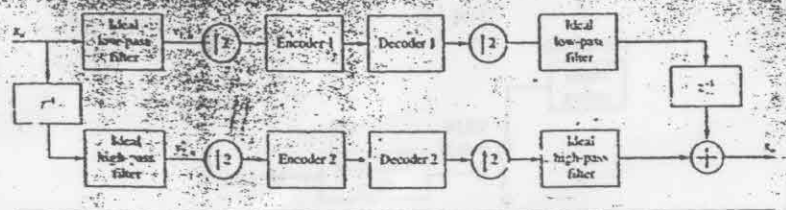


FIGURE 14. 11 Decomposition into two bands using ideal filters.

Let us look at how this system operates in the frequency domain. We begin by looking at the downsampling operation.

14.5.1 Downsampling ★

To see the effects of downsampling, we will obtain the Z-transform of the downsampled sequence in terms of the original source sequence. Because it is easier to understand what is going on if we can visualize the process, we will use the example of a source sequence that has the frequence profile shown in Figure 14.12. For this sequence the output of the ideal filters will have the shape shown in Figure 14.13.

Let's represent the downsampled sequence as $\{w_{1,n}\}$. The Z-transform $W_1(z)$ of the downsampled sequence $w_{1,n}$ is

$$W_1(z) = \sum w_{1,n} z^{-n}. \tag{14.24}$$

The downsampling operation means that

$$w_{1,n} = y_{1,2n}. \tag{14.25}$$

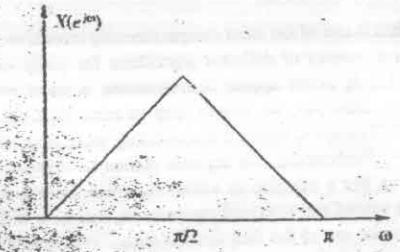


FIGURE 14. 12 Spectrum of the source output.

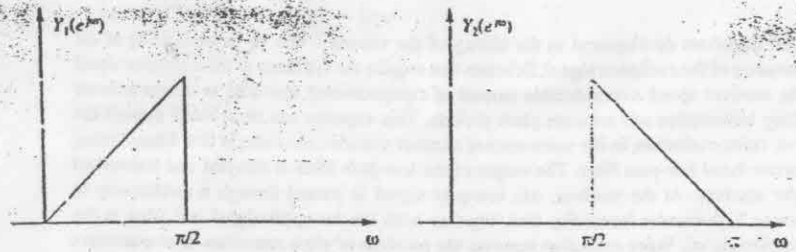


FIGURE 14. 13 Spectrum of the outputs of the ideal filters.

In order to find the Z-transform of this sequence, we go through a two-step process. Define the sequence

$$y'_{1,n} = \frac{1}{2}(1+e^{j\pi n})y_{1,n} \tag{14.26}$$

$$= \begin{cases} y_{1,n} & n \text{ even} \\ 0 & \text{otherwise.} \end{cases} \tag{14.27}$$

We could also have written Equation (14.26) as

$$y'_{1,n} = \frac{1}{2}(1+(-1)^n)y_{1,n}$$

however, writing the relationship as in Equation (14.26) makes it easier to extend this development to the case where we divide the source output into more than two bands.

The Z-transform of $y'_{1,n}$ is given as

$$Y'_1(z) = \sum_{n=-\infty}^{\infty} \frac{1}{2}(1+e^{j\pi n})y_{1,n}z^{-n}. \tag{14.28}$$

Assuming all summations converge,

$$Y'_1(z) = \frac{1}{2}\sum_{n=-\infty}^{\infty} y_{1,n}z^{-n} + \frac{1}{2}\sum_{n=-\infty}^{\infty} y_{1,n}(ze^{-j\pi})^{-n} \tag{14.29}$$

$$= \frac{1}{2}Y_1(z) + \frac{1}{2}Y_1(-z) \tag{14.30}$$

www.Vidyarthiplus.com
www.Vidyarthiplus.com

An important development in the history of the vocoders was an understanding of the importance of the excitation signal. Schemes that require the synthesis of the excitation signal at the receiver spend a considerable amount of computational resources to obtain accurate voicing information and accurate pitch periods. This expense can be avoided through the use of voice excitation. In the voice-excited channel vocoder, the voice is first filtered using a narrow-band low-pass filter. The output of the low-pass filter is sampled and transmitted to the receiver. At the receiver, this low-pass signal is passed through a nonlinearity to generate higher-order harmonics that, together with the low-pass signal, are used as the excitation signal. Voice excitation removes the problem of pitch extraction. It also removes the necessity for declaring every segment either voiced or unvoiced. As there are usually quite a few segments that are neither totally voiced or unvoiced, this can result in a substantial increase in quality. Unfortunately, the increase in quality is reflected in the high cost of transmitting the low-pass filtered speech signal.

The channel vocoder, although historically the first approach to analysis/synthesis—indeed the first approach to speech compression—is not as popular as some of the other schemes described here. However, all the different schemes can be viewed as descendants of the channel vocoder.

17.3.2 The Linear Predictive Coder (Government Standard LPC-10)

Of the many descendants of the channel vocoder, the most well-known is the linear predictive coder (LPC). Instead of the vocal tract being modeled by a bank of filters, in the linear predictive coder the vocal tract is modeled as a single linear filter whose output y_n is related to the input ϵ_n by

$$y_n = \sum_{i=1}^{M} b_i y_{n-i} + G\epsilon_n \qquad (17.1)$$

where G is called the gain of the filter. As in the case of the channel vocoder, the input to the vocal tract filter is either the output of a random noise generator or a periodic pulse generator. A block diagram of the LPC receiver is shown in Figure 17.5.

At the transmitter, a segment of speech is analyzed. The parameters obtained include a decision as to whether the segment of speech is voiced or unvoiced, the pitch period if the segment is declared voiced, and the parameters of the vocal tract filter. In this section, we will take a somewhat detailed look at the various components that make up the linear predictive coder. As an example, we will use the specifications for the 2.4-kbit U.S. Government Standard LPC-10.

The input speech is generally sampled at 8000 samples per second. In the LPC-10 standard, the speech is broken into 180 sample segments, corresponding to 22.5 milliseconds of speech per segment.

The Voiced/Unvoiced Decision

If we compare Figures 17.2 and 17.3, we can see there are two major differences. Notice that the samples of the voiced speech have larger amplitude; that is, there is more energy in

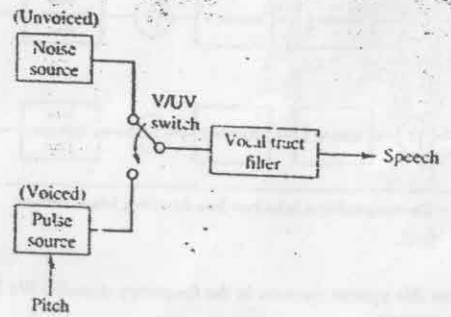


FIGURE 17.5 A model for speech synthesis.

the voiced speech. Also, the unvoiced speech contains higher frequencies. As both speech segments have average values close to zero, this means that the unvoiced speech waveform crosses the $x = 0$ line more often than the voiced speech sample. Therefore, we can get a fairly good idea about whether the speech is voiced or unvoiced based on the energy in the segment relative to background noise and the number of zero crossings within a specified window. In the LPC-10 algorithm, the speech segment is first low-pass filtered using a filter with a bandwidth of 1 kHz. The energy at the output relative to the background noise is used to obtain a tentative decision about whether the signal in the segment should be declared voiced or unvoiced. The estimate of the background noise is basically the energy in the unvoiced speech segments. This tentative decision is further refined by counting the number of zero crossings and checking the magnitude of the coefficients of the vocal tract filter. We will talk more about this latter point later in this section. Finally, it can be perceptually annoying to have a single voiced frame sandwiched between unvoiced frames. The voicing decision of the neighboring frames is considered in order to prevent this from happening.

Estimating the Pitch Period

Estimating the pitch period is one of the most computationally intensive steps of the analysis process. Over the years a number of different algorithms for pitch extraction have been developed. In Figure 17.2, it would appear that obtaining a good estimate of the pitch should be relatively easy. However, we should keep in mind that the segment shown in Figure 17.2 consists of 800 samples, which is considerably more than the samples available to the analysis algorithm. Furthermore, the segment shown here is noise-free and consists entirely of a voiced input. For a machine to extract the pitch from a short noisy segment which may contain both voiced and unvoiced components, can be a difficult undertaking.

Several algorithms make use of the fact that the autocorrelation of a periodic function $R_{xx}(k)$ will have a maximum when k is equal to the pitch period. Coupled with the fact that the estimation of the autocorrelation function generally leads to a smoothing out of the noise, this makes the autocorrelation function a useful tool for obtaining the pitch period.

www.Vidyarthiplus.com
www.Vidyarthiplus.com

Unfortunately, there are also some problems with the use of the autocorrelation. Voiced speech is not exactly periodic, which makes the maximum lower than we would expect from a periodic signal. Generally, a maximum is detected by checking the autocorrelation value against a threshold; if the value is greater than the threshold, a maximum is declared to have occurred. When there is uncertainty about the magnitude of the maximum value, it is difficult to select a value for the threshold. Another problem occurs because of the interference due to other resonances in the vocal tract. There are a number of algorithms that resolve these problems in different way. (see [105, 104] for details).

In this section, we will describe a closely related technique, employed in the LPC-10 algorithm, that uses the average magnitude difference function (AMDF). The AMDF is defined as

$$AMDF(P) = \frac{1}{N}\sum_{i=i_0+1}^{i_0+N} |y_i - y_{i-P}| \qquad (17.2)$$

If a sequence $\{y_n\}$ is periodic with period P_0, samples that are P_0 apart in the $\{y_n\}$ sequence will have values close to each other, and therefore the AMDF will have a minimum at P_0. If we evaluate this function using the /e/ and /s/ sequences, we get the results shown in Figures 17.6 and 17.7. Notice that not only do we have a minimum when P equals the pitch period, but any spurious minimums we may obtain in the unvoiced segments are very shallow; that is, the difference between the minimum and average values is quite small. Therefore, the AMDF can serve a dual purpose: it can be used to identify the pitch period as well as the voicing condition.

The job of pitch extraction is simplified by the fact that the pitch period in humans tends to fall in a limited range. Thus, we do not have to evaluate the AMDF for all possible values of P. For example, the LPC-10 algorithm assumes that the pitch period is between 2.5 and

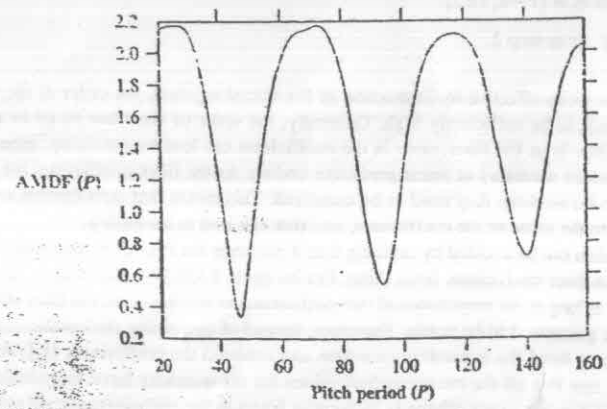


FIGURE 17.6 AMDF function for the sound /e/ in test.

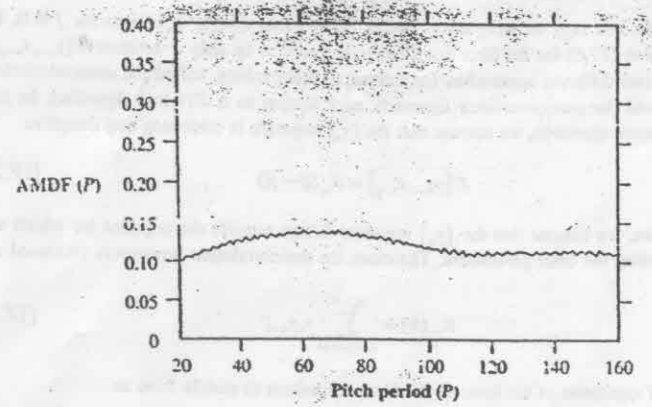


FIGURE 17.7 AMDF function for the sound /s/ in test.

19.5 milliseconds. Assuming a sampling rate of 8000 samples a second, this means that P is between 20 and 160.

Obtaining the Vocal Tract Filter

In linear predictive coding, the vocal tract is modeled by a linear filter with the input-output relationship shown in Equation (17.1). At the transmitter, during the analysis phase we obtain the filter coefficients that best match the segment being analyzed in a mean squared error sense. That is, if $\{y_n\}$ are the speech samples in that particular segment, then we want to choose $\{a_i\}$ to minimize the average value of e_n^2 where

$$e_n^2 = \left(y_n - \sum_{i=1}^{M} a_i y_{n-i} - G\epsilon_n\right)^2. \qquad (17.3)$$

If we take the derivative of the expected value of e_n^2 with respect to the coefficients $\{a_i\}$, we get a set of M equations:

$$\frac{\delta}{\delta a_j} E\left[\left(y_n - \sum_{i=1}^{M} a_i y_{n-i} - G\epsilon_n\right)^2\right] = 0 \qquad (17.4)$$

$$\Rightarrow -2E\left[\left(y_n - \sum_{i=1}^{M} a_i y_{n-i} - G\epsilon_n\right) y_{n-j}\right] = 0 \qquad (17.5)$$

$$\Rightarrow \sum_{i=1}^{M} a_i E[y_{n-i} y_{n-j}] = E[y_n y_{n-j}] \qquad (17.6)$$

www.Vidyarthiplus.com
www.Vidyarthiplus.com

where in the last step we have made use of the fact that $E[\epsilon_n y_{n-j}]$ is zero for $j \neq 0$. In order to solve (17.6) for the filter coefficients, we need to be able to estimate $E[y_{n-i}y_{n-j}]$. There are two different approaches for estimating these values, called the *autocorrelation* approach and the *autocovariance* approach, each leading to a different algorithm. In the autocorrelation approach, we assume that the $\{y_n\}$ sequence is stationary and therefore

$$E\left[y_{n-i}y_{n-j}\right] = R_{yy}(|i-j|) \tag{17.7}$$

Furthermore, we assume that the $\{y_n\}$ sequence is zero outside the segment for which we are calculating the filter parameters. Therefore, the autocorrelation function is estimated as

$$R_{yy}(k) = \sum_{n=n_0+1+k}^{n_0+N} y_n y_{n-k} \tag{17.8}$$

and the M equations of the form of (17.6) can be written in matrix form as

$$\mathbf{R}A = P \tag{17.9}$$

where

$$\mathbf{R} = \begin{bmatrix} R_{yy}(0) & R_{yy}(1) & R_{yy}(2) & \cdots & R_{yy}(M-1) \\ R_{yy}(1) & R_{yy}(0) & R_{yy}(1) & \cdots & R_{yy}(M-2) \\ R_{yy}(2) & R_{yy}(1) & R_{yy}(0) & \cdots & R_{yy}(M-3) \\ \vdots & \vdots & \vdots & & \vdots \\ R_{yy}(M-1) & R_{yy}(M-2) & R_{yy}(M-3) & \cdots & R_{yy}(0) \end{bmatrix} \tag{17.10}$$

$$A = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_M \end{bmatrix} \tag{17.11}$$

and

$$P = \begin{bmatrix} R_{yy}(1) \\ R_{yy}(2) \\ R_{yy}(3) \\ \vdots \\ R_{yy}(M) \end{bmatrix} \tag{17.12}$$

This matrix equation can be solved directly to find the filter coefficients

$$A = \mathbf{R}^{-1}P. \tag{17.13}$$

However, the special form of the matrix R obviates the need for computing $\mathbf{R}^{-1}$. Note that not only is R symmetric, but also each diagonal of R consists of the same element. For example, the main diagonal contains only the element $R_{yy}(0)$, while the diagonals above and below the main diagonal contain only the element $R_{yy}(1)$. This special type of matrix is called a *Toeplitz matrix*, and there are a number of efficient algorithms available for the inversion of Toeplitz matrices [224]. Because R is Toeplitz, we can obtain a recursive solution to (17.9) that is computationally very efficient and that has an added attractive feature from the point of view of compression. This algorithm is known as the Levinson-Durbin algorithm [225, 226]. We describe the algorithm without derivation. For details of the derivation, see [227, 105].

In order to compute the filter coefficients of an Mth-order filter, the Levinson-Durbin algorithm requires the computation of all filters of order less than M. Furthermore, during the computation of the filter coefficients, the algorithm generates a set of constants k_i known as the *reflection* coefficients, or *partial correlation* (PARCOR) coefficients. In the algorithm description below, we denote the order of the filter using superscripts. Thus, the coefficients of the fifth-order filter would be denoted by $\{a_i^{(5)}\}$. The algorithm also requires the computation of the estimate of the average error $E[\epsilon_n^2]$. We will denote the average error using an mth-order filter by E_m. The algorithm proceeds as follows:

1. Set $E_0 = R_{yy}(0)$, $i = 0$.
2. Increment i by one.
3. Calculate $k_i = \left(\sum_{j=1}^{i-1} a_j^{(i-1)} R_{yy}(i-j+1) - R_{yy}(i)\right) / E_{i-1}$.
4. Set $a_i^{(i)} = k_i$.
5. Calculate $a_j^{(i)} = a_j^{(i-1)} + k_i a_{i-j}^{(i-1)}$ for $j = 1, 2, \ldots, i-1$.
6. Calculate $E_i = (1 - k_i^2) E_{i-1}$.
7. If $i < M$, go to step 2.

In order to get an effective reconstruction of the voiced segment, the order of the vocal tract filter needs to be sufficiently high. Generally, the order of the filter is 10 or more. Because the filter is an IIR filter, error in the coefficients can lead to instability, especially for the high orders necessary in linear predictive coding. As the filter coefficients are to be transmitted to the receiver, they need to be quantized. This means that quantization error is introduced into the value of the coefficients, and that can lead to instability.

This problem can be avoided by noticing that if we know the PARCOR coefficients, we can obtain the filter coefficients from them. Furthermore, PARCOR coefficients have the property that as long as the magnitudes of the coefficients are less than one, the filter obtained from them is guaranteed to be stable. Therefore, instead of quantizing the coefficients $\{a_i\}$ and transmitting them, the transmitter quantizes and transmits the coefficients $\{k_i\}$. As long as we make sure that all the reconstruction values for the quantizer have magnitudes less than one, it is possible to use relatively high-order filters in the analysis/synthesis schemes.

www.Vidyarthiplus.com
www.Vidyarthiplus.com

The assumption of stationarity that was used to obtain (17.6) is not really valid for speech signals. If we discard this assumption, the equations to obtain the filter coefficients change. The term $E[y_{n-i}y_{n-j}]$ is now a function of both i and j. Defining

$$c_{ij} = E[y_{n-i}y_{n-j}] \tag{17.14}$$

we get the equation

$$CA = S \tag{17.15}$$

where

$$C = \begin{bmatrix} c_{11} & c_{12} & c_{13} & \cdots & c_{1M} \\ c_{21} & c_{22} & c_{23} & \cdots & c_{2M} \\ \vdots & \vdots & \vdots & & \vdots \\ c_{M1} & c_{M2} & c_{M3} & \cdots & c_{MM} \end{bmatrix} \tag{17.16}$$

and

$$S = \begin{bmatrix} c_{10} \\ c_{20} \\ c_{30} \\ \vdots \\ c_{M0} \end{bmatrix} \tag{17.17}$$

The elements c_{ij} are estimated as

$$c_{ij} = \sum_{n=n_0+1}^{n_0+N} y_{n-i}y_{n-j}. \tag{17.18}$$

Notice that we no longer assume that the values of y_n outside of the segment under consideration are zero. This means that in calculating the C matrix for a particular segment, we use samples from previous segments. This method of computing the filter coefficients is called the *covariance method*.

The C matrix is symmetric but no longer Toeplitz, so we can't use the Levinson-Durbin recursion to solve for the filter coefficients. The equations are generally solved using a technique called the *Cholesky decomposition*. We will not describe the solution technique here. (You can find it in most texts on numerical techniques; an especially good source is [178].) For an in-depth study of the relationship between the Cholesky decomposition and the reflection coefficients, see [228].

The LPC-10 algorithm uses the covariance method to obtain the reflection coefficients. It also uses the PARCOR coefficients to update the voicing decision. In general, for voiced signals the first two PARCOR coefficients have values close to one. Therefore, if both the first two PARCOR coefficients have very small values, the algorithm sets the voicing decision to unvoiced.

Transmitting the Parameters

Once the various parameters have been obtained, they need to be coded and transmitted to the receiver. There are a variety of ways this can be done. Let us look at how the LPC-10 algorithm handles this task.

The parameters that need to be transmitted include the voicing decision, the pitch period, and the vocal tract filter parameters. One bit suffices to transmit the voicing information. The pitch is quantized to 1 of 60 different values using a log-companded quantizer. The LPC-10 algorithm uses a 10th-order filter for voiced speech and a 4th-order filter for unvoiced speech. Thus, we have to send 11 values (10 reflection coefficients and the gain) for voiced speech and 5 for unvoiced speech.

The vocal tract filter is especially sensitive to errors in reflection coefficients that have magnitudes close to one. As the first few coefficients are most likely to have values close to one, the LPC-10 algorithm specifies the use of nonuniform quantization for k_1 and k_2. The nonuniform quantization is implemented by first generating the coefficients

$$g_i = \frac{1+k_i}{1-k_i} \tag{17.19}$$

which are then quantized using a 5-bit uniform quantizer. The coefficients k_3 and k_4 are both quantized using a 5-bit uniform quantizer. In the voiced segments, coefficients k_5 through k_8 are quantized using a 4-bit uniform quantizer, k_9 is quantized using a 3-bit uniform quantizer, and k_{10} is quantized using a 2-bit uniform quantizer. In the unvoiced segments, the 21 bits used to quantize k_5 through k_{10} in the voiced segments are used for error protection.

The gain G is obtained by finding the root mean squared (rms) value of the segment and quantized using 5-bit log-companded quantization. Including an additional bit for synchronization, we end up with a total of 54 bits per frame. Multiplying this by the total number of frames per second gives us the target rate of 2400 bits per second.

Synthesis

At the receiver, the voiced frames are generated by exciting the received vocal tract filter by a locally stored waveform. This waveform is 40 samples long. It is truncated or padded with zeros depending on the pitch period. If the frame is unvoiced, the vocal tract is excited by a pseudorandom number generator.

The LPC-10 coder provides intelligible reproduction at 2.4 kbits. The use of only two kinds of excitation signals gives an artificial quality to the voice. This approach also suffers when used in noisy environments. The encoder can be fooled into declaring segments of speech unvoiced because of background noise. When this happens, the speech information gets lost.

17.3.3 Code Excited Linear Predicton (CELP)

As we mentioned earlier, one of the most important factors in generating natural-sounding speech is the excitation signal. As the human ear is especially sensitive to pitch errors, a great deal of effort has been devoted to the development of accurate pitch detection algorithms.

www.Vidyarthiplus.com
www.Vidyarthiplus.com

www.Vidyarthiplus.com

PART-A

1) Write the equation for Blackman Window?

Solu:

The window function of a causal Blackman window is expressed by

$$W_B(n) = \begin{cases} 0.42 - 0.5\cos\frac{2\pi n}{M-1} + 0.08\cos\frac{4\pi n}{M-1} & 0 \le n \le M-1 \\ 0 & \text{otherwise} \end{cases}$$

The window function of a non causal Blackman window is given by

$$W_B(n) = \begin{cases} 0.42 + 0.5\cos\frac{2\pi n}{M-1} + 0.08\frac{4\pi n}{M-1} & \text{for } |n| < \frac{M-1}{2} \\ 0 & \text{otherwise} \end{cases}$$

2) What is zero input limit cycle oscillation?

Sol:

When a stable IIR digital filter is excited by finite input sequence i.e constant, the output will ideally decay to zero. The non-linearities due to finite precision arithmetic operation often cause periodic oscillation to occur in the output. Such oscillation is recursive system are called zero input limit cycle oscillation.

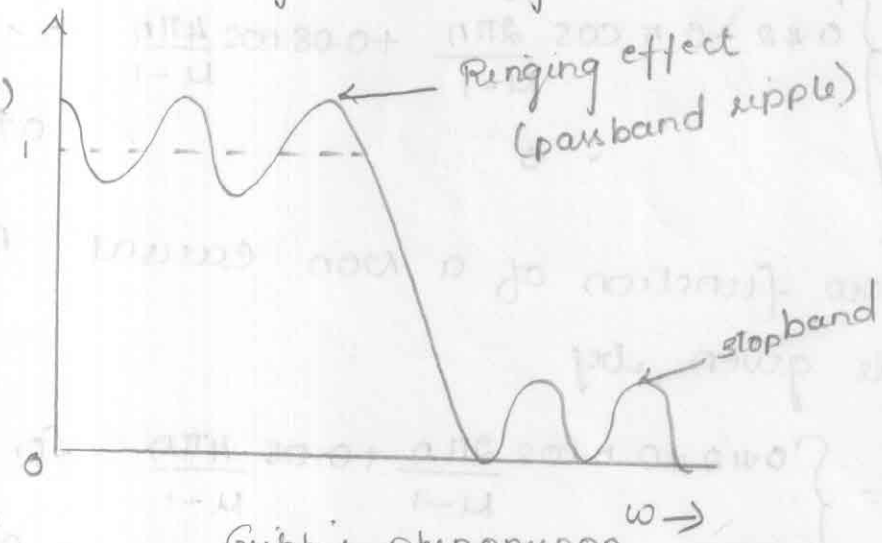
3) What is Gibb's phenomenon?

Sol: In the design of FIR filter the infinite duration impulse response is truncated in the

www.Vidyarthiplus.com

www.Vidyarthiplus.com

finite duration impulse response. The abrupt truncation introduced oscillations in pass band and stop band. This effect know as Gibb's phenomenon or Gibb's oscillations. These oscillation cannot be avoided fully but can be reduced to some extent by choosing the appropriate window



Gibb's phenomenon

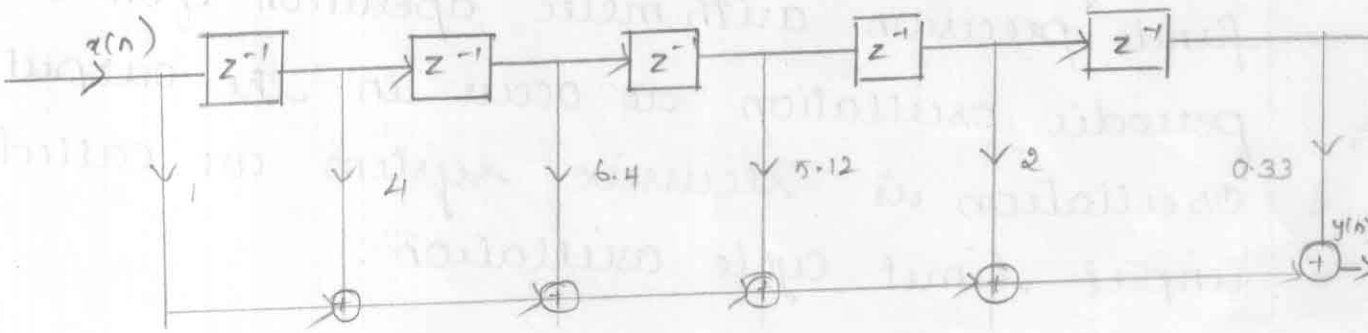
4) Realize the FIR transfer function in direct form

$$H(z) = 1+4z^{-1}+6.4z^{-2}+5.12z^{-3}+2z^{-4}+0.33z^{-5}$$

Sol:

$$H(z) = \frac{Y(z)}{X(z)} = 1+4z^{-1}+6.4z^{-2}+5.12z^{-3}+2z^{-4}+0.33z^{-5}$$

$$H(z) \Rightarrow Y(z) = 1X(z)+4z^{-1}X(z)+6.4z^{-2}X(z)+5.12z^{-3}X(z)+2z^{-4}X(z)+0.33z^{-5}X(z)$$



5) Define phase delay?

Sol: phase delay is defined by $\tau_p = -\frac{\phi(\omega)}{\omega}$

www.Vidyarthiplus.com

www.Vidyarthiplus.com

$$\phi(\omega) = -\tau\omega$$

$$\tau = \frac{M-1}{2}$$

6) Define truncation error for sign magnitude representation and for 2's complement.

Sol.:

Truncation error for sign magnitude:

Truncation error is negative and the range is

$$-(2^{-B} - 2^{-L}) \le E_T < 0$$

Truncation error is positive

$$0 \le E_T \le (2^{-B} - 2^{-L})$$

The overall range of the truncation error for the sign magnitude representation

$$-(2^{-B} - 2^{-L}) \le E_T \le (2^{-B} - 2^{-L})$$

Truncation error for two's complement representation:

$$-(2^{-B} - 2^{-L}) \le E_T \le 0$$

7) What are the types of limit cycle oscillation?

Sol.:

(i) zero input limit cycle oscillation

(ii) overflow limit cycle oscillation

8) What is overflow oscillations?

Sol.: In addition to limit cycle oscillation caused by rounding the result of multiplication, there are several type limit oscillation caused by addition which makes the filter output oscillation between

www.Vidyarthiplus.com

www.Vidyarthiplus.com

maximum and minimum amplitudes. Such limit cycle is called over flow limit cycle oscillation.

9) What are the advantage of floating point arithmetic?

Sol:

→ Larger dynamic range.

→ overflow in floating point representation is unlikely.

10) What is the major disadvantage of window design method?

Sol:

→ The designing of FIR filter using window is not flexible.

→ In many application the expression for the desired filter response is complicated.

11) Why co-efficient quantization effect are not sever in linear phase FIR filter?

Sol: The filter coefficients are computed to infinite precision of theory. If quantized the the frequency response of resulting filter may differ from desired response and sometimes the filter may fail to meet the desired specification. If the poles of desired filter are close to unit circle. Thus the filter with quantized coefficients lie outside unit circle leading to instability. The error that arise due to this instability is

www.Vidyarthiplus.com

www.Vidyarthiplus.com

called coefficient error.

"Thus the coefficient quantization error is designed for infinite impulse response so it does not serve FIR linear phase FIR filter".

12) Represent decimal number 0.69 in fixed point representation of length $N=6$.

Sol:

Decimal number $= 0.69$

$N = 6$

Binary form $= 0.101100$

One's complement $= 1.010011$

2's complement $= 0.000001$

1.010100

$$(0.69)_{10} = (1.010100)_2$$

13) What are limit cycle:

Sol: The limit cycle occur as a result of the quantization effects in multiplication. The amplitude of the output during a limit cycle are confined to a range of values that is called deadband of the filter.

$$y(n) = \alpha y(n-1) + x(n), n > 0$$

14) Compare FIR filter and FIR filter with regards

(a) Stability:

www.Vidyarthiplus.com

www.Vidyarthiplus.com

FIR filters are inherently stable :

We know that LTI system is said to be stable if bounded input produce bounded output (BIBO)

$$y(n) = \sum_{k=0}^{M-1} b_k x(n-k)$$

$$y(n) = h(0)x(n) + h(1)x(n-1) + \dots + h(M-1)x(N-M+1)$$

here $h(0), h(1) \dots h(M-1)$ are constant that means they are bounded, the output will be bounded if we apply bounded input. That means for every bounded input, the output of FIR filter is bounded. Thus FIR filters are inherently stable.

www.Vidyarthiplus.com

www.Vidyarthiplus.com

PART-B

1) Compare Hamming window and Kaiser window.

Hamming	Kaiser
1) The main lob width is equal to $\frac{8\pi}{N}$ and the peak lobe level is -41 dB	The main lobe with the peak side lobe level can be varied by varying the parameter α and N.
2) The low pass FIR filter designed will have first side lobe peak of -53dB	The side lobe peak can be varied by varying the parameter α.

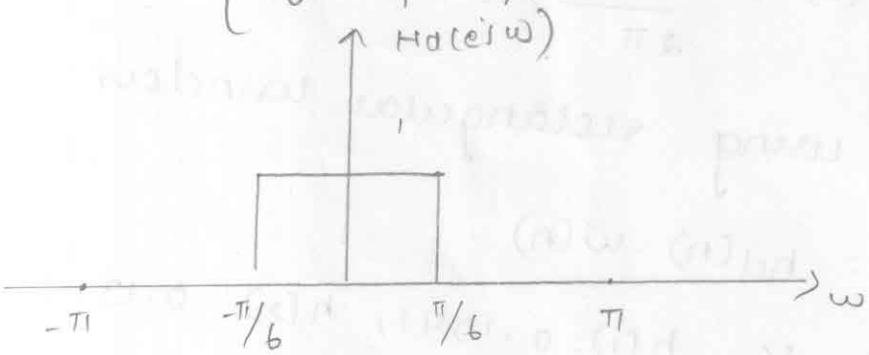
2) For a FIR linear phase digital filter approximates the ideal frequency response.

$$H_d(\omega) = 1 \quad |\omega| \le \pi/6$$
$$\quad\;\; 0 \quad \pi/6 \le |\omega| \le \pi$$

Determine the coefficient of a 5 tap filter using rectangular window.

Sol:

$$H_d(e^{j\omega}) = \begin{cases} 1 & \text{for } |\omega| \le \pi/6 \\ 0 & \text{for } \pi/6 \le |\omega| \le \pi \end{cases}$$



$$h(n) = \frac{1}{2\pi}\int_{-\pi}^{\pi} H_d(e^{j\omega})\, e^{j\omega n}\, d\omega$$

www.Vidyarthiplus.com

www.Vidyarthiplus.com

$$= \frac{1}{2\pi}\int_{-\pi/6}^{\pi/6} (1) e^{j\omega n}\, d\omega \quad = \frac{1}{2\pi}\left[\frac{e^{j\omega n}}{jn}\right]$$

$$= \left[\frac{1}{2\pi jn} e^{j\omega n}\right]_{-\pi/6}^{\pi/6}$$

$$= \frac{1}{2\pi jn}\left[e^{j\pi/6 n} - e^{-j\pi/6 n}\right]$$

$$h_d(n) = \frac{\sin\frac{\pi}{6}n}{\pi n} \quad |n| \le \frac{M-1}{2} \le \frac{5-1}{2} \le \frac{4}{2} \le 2$$

$$= \frac{\sin\frac{\pi}{6}n}{\pi n} \quad |n| \le 2$$

for $n=0$, substitute in eq (1) we get

$$h_d(n) = \frac{1}{2\pi}\int_{-\pi/6}^{\pi/6} e^{-j\omega(0)}\, d\omega$$

$$h_d(0) = \frac{1}{6}$$

$$n=1, \quad h_d(1) = \frac{\sin \pi/6}{\pi} = 0.1591$$

$$n=2, \quad h_d(2) = \frac{\sin \pi/3}{2\pi} = 0.1378$$

(a) By using rectangular window

$$h(n) = h_d(n) \cdot w(n)$$

$\therefore h(0) = 1/6$, $h(1) = 0.1591$, $h(2) = 0.1378$

The frequency response is given by

www.Vidyarthiplus.com

www.Vidyarthiplus.com

$$H(e^{j\omega}) = \sum_{n=0}^{2} h(n) e^{-j\omega n}$$

$$= h(0)e^{0} + h(1)e^{-j\omega} + h(2)e^{-j2\omega}$$

$$= 1/6 + 0.1591 e^{-j\omega} + 0.1378 e^{-j2\omega}$$

3) Given a three stages lattice filter with coefficients $k_1 = \frac{1}{4}$, $k_2 = \frac{1}{4}$, $k_3 = 1/4$ determine the FIR filter coefficient for the direct form structure.

Sol:

$\alpha_3(0) = 1$; $\alpha_3(3) = k_3 = 1/4$

$\alpha_1(1) = k_1 = \frac{1}{2}$, $\alpha_2(2) = k_2 = \frac{1}{3}$

$$\alpha_m(k) = \alpha_{m-1}(k) + k_m \alpha_{m-1}(m-k)$$

for $m = 2$ and $k = 1$

$$\alpha_2(1) = \alpha_1(1) + k_2 \alpha_1(1)$$

$$= \frac{1}{2} + \frac{1}{3} \cdot \frac{1}{2} = \frac{3+1}{6} = \frac{2}{3}$$

$m = 3$ + $k = 1$

$$\alpha_3(1) = \alpha_2(1) + \alpha_3(3)\alpha_2(2)$$

$$= \frac{2}{3} + \frac{1}{4} \cdot \frac{1}{3} = \frac{8+1}{12} = 3/4$$

$m = 2$ and $k = 2$

$$\alpha_3(2) = \alpha_2(2) + \alpha_3(3)\alpha_2(1)$$

$$= \frac{1}{3} + \frac{1}{4} \cdot \frac{2}{3} = \frac{4+2}{12} = 1/2$$

$\alpha_3(0) = 1$; $\alpha_3(1) = 3/4$, $\alpha_3(2) = 1/2$; $\alpha_3(3) = 1/4$

www.Vidyarthiplus.com

www.Vidyarthiplus.com

4) Determine the coefficient of a linear phase FIR filter of length $M=15$ has a symmetric unit samples response and a frequency response that satisfies the condition b

$$H\left(\frac{2\pi k}{15}\right) = \begin{cases} 1 & k=0,1,2,3 \\ 0 & k=4,5,6,7 \end{cases}$$

Sol:

$$H(k) = H_d(e^{j\omega}) \left[\omega = \frac{2\pi k}{M}\right]$$

$M = 15$

$$\therefore H(k) = e^{-j\omega\left(\frac{M-1}{2}\right)}$$

$$= e^{-j\frac{2\pi k}{15}\left(\frac{15-1}{2}\right)}$$

$$H(k) = e^{-j14\pi k/15} \quad k=0,1 \ldots M-1$$

$\omega_k = \frac{2\pi k}{15}$ for $0 \le \omega \le \pi/2$ the value of $k=0,1,2,3$

for $\pi/2 \le \omega \le 3\pi/2$, $k=2,3,4,5,6,7,8,9,10,11$ and

for $\frac{3\pi}{2} \le \omega \le 2\pi$, $k=12,13,14$

$$H(k) = \begin{cases} 1 & \text{for } 0 \le k \le 3 \text{ and } 12 \le k \le 14 \\ 0 & \text{for } 4 \le k \le 11 \end{cases}$$

M odd :

$$H(k) = \begin{cases} e^{-j14\pi k/15} & \text{for } k=0,1,2,3 \\ 0 & \text{for } k=4,5,6,7,8,9,10 \\ e^{-j4\pi(k-15)/15} & \text{for } k=11,12,13,14 \end{cases}$$

$$\therefore h(n) = \frac{1}{M}\left[H(0) + 2\sum_{k=1}^{\frac{M-1}{2}} Re\left[(H(k)e^{j2\pi nk/15}\right]\right]$$

www.Vidyarthiplus.com

www.Vidyarthiplus.com

$$= \frac{1}{15}\left[1 + 2\sum_{k=1}^{7} \mathrm{Re}\left(e^{-j14\pi k/15} \cdot e^{-j2\pi nk/15}\right)\right]$$

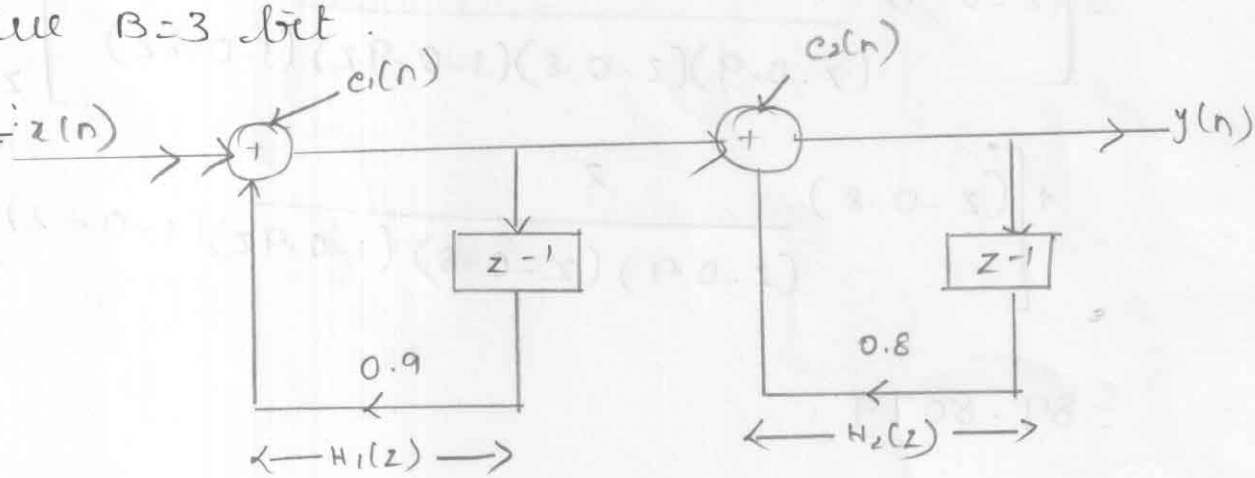
$$= \frac{1}{15}\left[1 + 2\sum_{k=1}^{3} \cos\frac{2\pi k(7-n)}{15}\right]$$

$$h(n) = \frac{1}{15}\left[1 + 2\cos\frac{2\pi(7-n)}{15} + 2\cos\frac{4\pi(7-n)}{15} + 2\cos\frac{6\pi(7-n)}{15}\right]$$

$$n = 0, 1, \ldots 14$$

5) For the given transfer function $H(z) = H_1(z) \cdot H_2(z)$ Where $H_1(z) = \frac{1}{1-0.9z^{-1}}$ and $H_2(z) = \frac{1}{1-0.8z^{-1}}$ Find the output round off noise power calculate the value B=3 bit.

Sol: $x(n)$



$$H(z) = H_1(z) \cdot H_2(z)$$

$$= \frac{1}{(1-0.9z^{-1})(1-0.8z^{-1})}$$

Overall output round off noise power is given by $\sigma_{ov}^2 = \sigma_{01}^2 + \sigma_{02}^2$

$$\sigma_{01}^2 = \frac{\sigma_e^2}{2\pi j}\oint_c H(z)\,H(z^{-1})\,z^{-1}\,dz$$

www.Vidyarthiplus.com

www.Vidyarthiplus.com

$$= \frac{\sigma_e^2}{2\pi j} \oint_c \frac{1}{(1-0.9z^{-1})(1-0.8z^{-1})} \times \frac{1}{(1-0.9z^{-1})(1-0.8z^{-1})} z^{-1} dz$$

$$= \sigma_e^2 \times I_1$$

$$I_1 = \frac{1}{2\pi j} \oint_c \frac{z}{(z-0.9)(z-0.8)(1-0.9z)(1-0.8z)} dz$$

By using residue method

I_1 = sum of the residues at the pole within the unit circle

= residue at $z=0.9$ + residue at $z=0.8$

$$= \left[(z-0.9) \frac{z}{(z-0.9)(z-0.8)(1-0.9z)(1-0.8z)}\right]_{z=0.9}$$

$$+ \left[(z-0.8) \frac{z}{(z-0.9)(z-0.8)(1-0.9z)(1-0.8z)}\right]_{z=0.8}$$

$$= 89.8079$$

$$\sigma_{01}^2 = \sigma_e^2 \times I_1$$
$$= 89.8079\, \sigma_e^2$$

$$\sigma_{02}^2 = \frac{\sigma_e^2}{2\pi j} \oint_c H_2(z) H_2(z^{-1}) z^{-1} dz$$

$$= \frac{\sigma_e^2}{2\pi j} \oint_c \frac{1}{(1-0.8z^{-1})} \frac{1}{(1-0.8z)} \times z^{-1} dz$$

$$= \frac{\sigma_e^2}{2\pi j} \oint_c \frac{dz}{(z-0.8)(1-0.8z)}$$

www.Vidyarthiplus.com

www.Vidyarthiplus.com

$$= \sigma_e^2 \times I_2$$

$$I_2 = \frac{1}{2\pi j} \oint_c \frac{dz}{(z-0.8)(1-0.8z)}$$

I_2 = sum of the residue at the pole within the unit circle

$$= \left[(z-0.8) \frac{1}{(z-0.8)(1-0.8z)} \right]_{z=0.8}$$

$$= 2.778$$

$$\sigma_{02}^2 = \sigma_e^2 \times I_2$$

$$= 2.778\, \sigma_e^2$$

The overall output round off noise power is

$$\sigma_{0r}^2 = \sigma_{01}^2 + \sigma_{02}^2$$

$$= \sigma_e^2\, 89.8079 + 2.778\, \sigma_e^2$$

$$= \sigma_e^2\, 92.586$$

$$\sigma_e^2 = \frac{2^{-2B}}{12}$$

$$\sigma_{0r}^2 = 92.586 \times \frac{2^{-2B}}{12}$$

$$= 92.586 \times \frac{2^{-6}}{12}$$

6) Distinguish between fixed point and floating point arithmetic.

Sol:

Fixed point	floating point
1) Fast operation	slow operation

www.Vidyarthiplus.com

www.Vidyarthiplus.com

$= \sigma_e^2 \times I_2$

$$I_2 = \frac{1}{2\pi j} \oint_c \frac{dz}{(z-0.8)(1-0.8z)}$$

I_2 = sum of the residue at the pole within the unit circle

$$= \left[(z-0.8) \frac{1}{(z-0.8)(1-0.8z)} \right]_{z=0.8}$$

$= 2.778$

$\sigma_{02}^2 = \sigma_e^2 \times I_2$

$= 2.778\, \sigma_e^2$

The overall output round off noise power is

$\sigma_{0v}^2 = \sigma_{01}^2 + \sigma_{02}^2$

$= \sigma_e^2\, 89.8079 + 2.778\, \sigma_e^2$

$= \sigma_e^2\, 92.586$

$\sigma_e^2 = \frac{2^{-2B}}{12}$

$\sigma_{0v}^2 = 92.586 \times \frac{2^{-2B}}{12}$

$= 92.586 \times \frac{2^{-6}}{12}$

6) Distinguish between fixed point and floating point arithmetic.

Sol:

Fixed point	floating point
1) Fast operation	slow operation

www.Vidyarthiplus.com

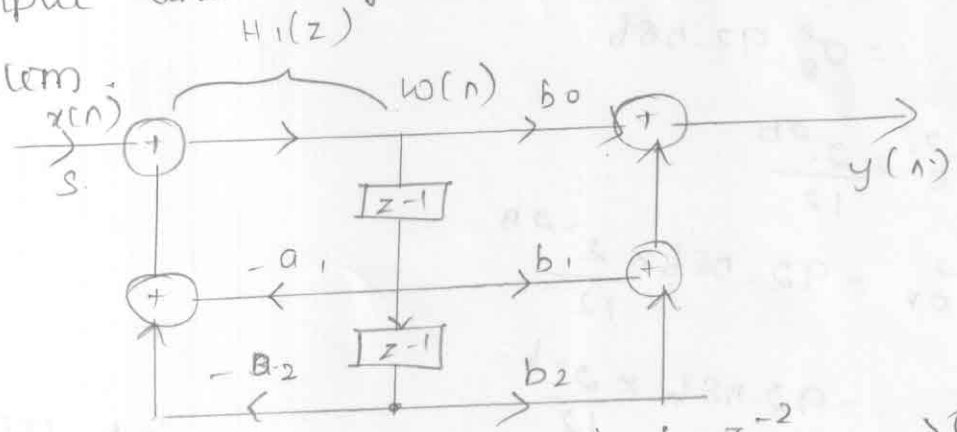
www.Vidyarthiplus.com

2) Round off error are present only in addition operation	Round off error occur in both addition and multiplication operation
3) Dynamic range is small	Dynamic range is large
4) overflow error is present	Overflow error is not present
5) Application are only in small computers.	Application are in large general purpose computer

7) With respect to finite word length effects in digital filter, with examples discuss about

(i) Signal scaling :

Due to overflow error, limit cycles gets eliminated which results into undesirable signal distortion. To avoid this effect it is required to scale the input signal and limit sample response between the input and any internal summing node in the system.



$$H(z) = \frac{Y(z)}{X(z)} = S\,\frac{b_0 + b_1 z^{-1} + b_2 z^{-2}}{1 + a_1 z^{-1} + a_2 z^{-2}} \longrightarrow (1)$$

$$\therefore H_1(z) = S\,\frac{N(z)}{D(z)} \longrightarrow (2)$$

Now consider transfer function for input signal and first internal summing node — (3)

$$H_1(z) = \frac{W(z)}{X(z)} = S\,\frac{1}{1 + a_1 z^{-1} + a_2 z^{-2}} = S\,\frac{1}{D(z)} \quad (4)$$

www.Vidyarthiplus.com

www.Vidyarthiplus.com

$$W(z) = H_1(z)\, X(z)$$

$$W(z) = S \frac{1}{D(z)} X(z)$$

$$= S \frac{X(z)}{D(z)}$$

$$\frac{1}{D(z)} = R(z)$$

$$\therefore W(z) = S\, X(z)\, R(z) \longrightarrow (5)$$

Taking inverse z-transform

$$w(n) = S \frac{1}{2\pi} \int X(e^{jQ}) \cdot R(e^{jQ})\, d\theta$$

squaring on both side

$$W^2(n) = \frac{S^2}{4\pi^2} \int \left[X[e^{jQ}]\, R(e^{jQ})\, dQ \right]^2$$

$$W^2(n) \le S^2 \left[\frac{1}{2\pi} \int |X(e^{jQ})|^2 dQ \right] \left[\frac{1}{2\pi} \int |R(e^{jQ})|^2 dQ \right.$$

using parseval's theorem

$$W^2(n) \le S^2 \sum_{n=0}^{\infty} h^2(n) \cdot \frac{1}{2\pi} \int |R(e^{jQ})|^2 dQ \quad \text{—} (6)$$

$$z = e^{jQ} \quad \text{—} (7)$$

$$dz = j e^{jQ} dQ = jz\, dQ$$

$$\Rightarrow dQ = \frac{dz}{jz} = \frac{z^{-1} dz}{j} \quad \text{—} (8)$$

$W^2(n)$ Sub eq (7) & (8) in (6)

$$W^2(n) \le S^2 \sum_{n=0}^{\infty} h^2(n) \cdot \frac{1}{2\pi j} \int |R(z)|^2 z^{-1} dz \quad \text{—} (9)$$

$$\frac{S^2}{2\pi j} \int |R(z)|^2 z^{-1} dz = 1 \quad \text{—} (10)$$

Eq (9) becomes

$$W^2(n) \le \sum_{n=0}^{\infty} x^2(n)$$

From eq (10)

www.Vidyarthiplus.com

www.Vidyarthiplus.com

$$S^2 = \frac{1}{\frac{1}{2\pi j}} \int |R(z)|^2 z^{-1} dz$$

$$= \frac{1}{\frac{1}{2\pi j} \int \frac{1}{D(z).D(z^{-1})} z^{-1} dz}$$

$$\therefore S^2 = \frac{1}{x}$$

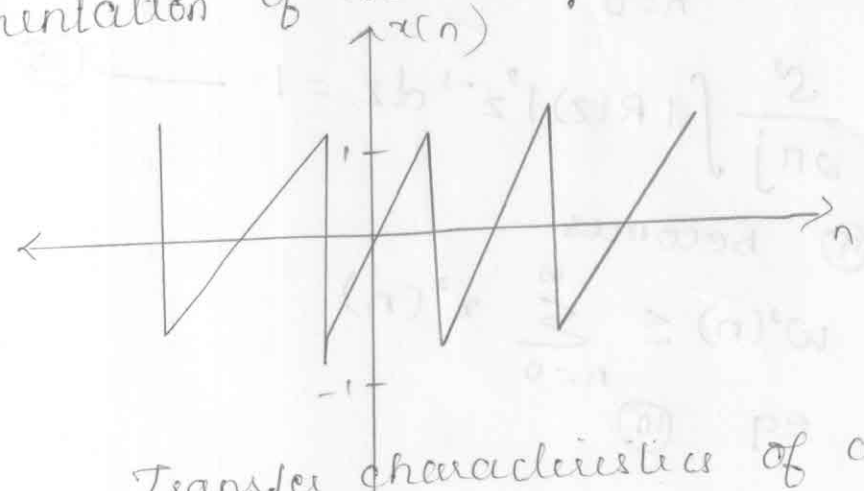
$$x = \frac{1}{2\pi j} \int \frac{1}{D(z).D(z^{-1})} z^{-1} dz$$

$$\therefore S = \frac{1}{\sqrt{x}}$$

This is the scaling factor used to scale the signal to overcome overflow errors.

ii) Overflow limit cycle oscillations:

The limit cycle oscillation caused by binary arithmetic, which make the fullu output oscillate between maximum and minimum amplitudes. Such limit cycle have been referred to as overflow oscillation. An overflow in addition of two or more binary number occurs when the sum exceeds the word size available in the digital implementation of the system.



Transfer characteristics of an adder

www.Vidyarthiplus.com

www.Vidyarthiplus.com

Let us consider two positive number n_1 & n_2

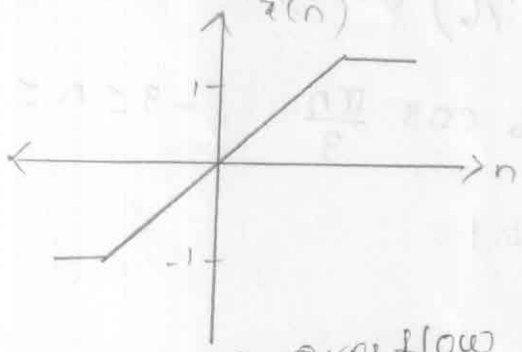
$n_1 = 0.110 \longrightarrow \frac{6}{8}$

$n_2 = 0.101 \longrightarrow \frac{5}{8}$

$n_1 + n_2 = 1.011 \longrightarrow -3/8$ in sign magnitude

Here, when two +ve number are added the sum is wrongly interpreted as a -ve number

The transfer characteristic of an adder is shown when n is the input to the adder and x(n) is the related output.



We can find that overflow occur if the total input is out of range (-1, 1).

8) Design a linear phase FIR digital filter for the given specification using hamming window of length $M = 7$

$$H_d(\omega) = \begin{cases} e^{-j3\omega} & \text{for } |\omega| < \pi/6 \\ 0 & \text{for } \pi/6 < |\omega| \le \pi \end{cases}$$

Sol:

Hamming window sequence is given by

$$W_{HM}(n) = 0.5 + 0.5 \cos \frac{2\pi}{N-1} \quad \text{for } |n| \le \frac{N-1}{2}$$

$$= 0 \quad \text{otherwise}$$

$$h_d(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} H(e^{j\omega}) e^{j\omega n} d\omega$$

www.Vidyarthiplus.com

www.Vidyarthiplus.com

$$= \frac{1}{2\pi} \int_{-\pi/6}^{\pi/6} e^{-j3\omega} e^{j\omega n} d\omega = \frac{1}{2\pi} \int_{-\pi/6}^{\pi/6} e^{j\omega(n-3)} d\omega$$

$$= \frac{1}{2\pi} \left[\frac{e^{j\omega(n-3)}}{j(n-3)} \right]_{-\pi/6}^{\pi/6}$$

$$= \frac{1}{\pi(n-3)} \left[\frac{e^{j\omega(n-3)}}{2j} \right]_{-\pi/6}^{\pi/6}$$

$$= \frac{1}{\pi(n-3)} \sin\left(\pi/6\right)(n-3)$$

$$w_H(n) = 0.54 + 0.46 \cos \frac{\pi n}{3}, \quad -3 \le n \le 3$$

$w_H(0) = 0.54 + 0.46 = 1$

$w_H(1) = 0.77$

$w_H(2) = 0.31$

$w_H(3) = 0.08$ Fourier coefficient

$$h_d(0) = h_d(6) = \frac{\sin \pi/6 \times 3}{\pi \times 3} = \frac{1}{9.42} = 0.1061$$

180°, 3.14

$$h_d(1) = h_d(5) = \frac{\sin \pi/6 \times 2}{\pi \times 2} = \frac{0.8660}{6.28} = 0.1379$$

$$h_d(2) = h_d(4) = \frac{\sin \pi/6 \times 1}{\pi \times 1} = \frac{0.5}{3.14} = 0.1592$$

$$h_d(3) = \lim_{n \to 0} \frac{\sin\left(\pi/6\right)(n-3)}{\pi(n-3)} = 1/6 = 0.166$$

www.Vidyarthiplus.com

www.Vidyarthiplus.com

The filter coefficient using hamming window

$$h(n) = \begin{cases} h_d(n) \cdot w_H(n) & \text{for } -3 \le n \le 3 \end{cases}$$

$h(0) = 0.1061 \times 1 = 0.1061$

$h(1) = 0.77 \times 0.1379 = 0.1062$

$h(2) = 0.1592 \times 0.31 = 0.0493$

$h(3) = 0.1666 \times 0.08 = 0.0133$

The transfer function of the filter is given by

$$H(z) = h(0) + \sum_{n=1}^{3} [h(n)(z^{-n} + z^{n})]$$

$$= 0.1061 + 0.1062(z^{-1} + z^{1}) + 0.0493(z^{-2} + z^{2}) + 0.0133(z^{-3} + z^{3})$$

The transfer fun of realizable filter

$$H'(z) = z^{-3} H(z)$$

$$= 0.1061 z^{-3} + 0.1062 z^{-4} + 0.1061 z^{-2} + 0.0493 z^{-5} + 0.0493 z^{-1} + 0.0133 z^{-6} + 0.0133$$

Filter coefficient of causal filter

$h(0) = h(6) = 0.0133$

$h(1) = h(5) = 0.0493$

$h(2) = h(4) = 0.1062$

$h(3) = 0.1061$

9) Explain the characteristics of a limit cycle oscillation with respect to the system described by the equation $y(n) = 0.9n\, y(n-1) + x(n)$. Determine the deadband of the filter $x(n) = (3/4)\,\delta(n)$. assume register length is 4

www.Vidyarthiplus.com

$\alpha = 0.95$

n	$x(n)$	$y(n-1)$	$\alpha y(n-1)$	$Q[0.95\, y(n-1)]$	$y_q(n) = Q[0.95\, y(n-1)] + x(n)$	
0	$x(0) = 0.75$	$y(-1) = 0$	$(0.0000)_2$	$(0.0000)_2 = (0)_{10}$	$y(0) = 0 + 0.75$ $= (0.1100)_2$ $= \frac{12}{16} = 3/4$	
1	$x(1) = 0$	$y(0) = 3/4$	$(0.7125)_{16}$ $= (0.10110)_2$	$(0.1011) = 11/16$	$y(1) = 11/16$	$0.7125 \times 2 = 1.425 \rightarrow 0$ $0.425 \times 2 = 0.85 \rightarrow 0$ $0.85 \times 2 = 1.7 \rightarrow 1$ $0.7 \times 2 = 1.4 \rightarrow 1$ $0.4 \times 2 = 0.8 \rightarrow 0$
2	$x(2) = 0$	$y(1) = 11/16$	$(0.653)_{16}$ $= (0.10100)_2$	$(0.1010)_2 = 5/8$	$y(2) = 5/8$	$0.653 \times 2 = 1.306 \rightarrow 1$ $0.306 \times 2 = 0.612 \rightarrow 0$ $0.612 \times 2 = 1.224 \rightarrow 1$ $0.224 \times 2 = 0.488 \rightarrow 0$ $0.488 \times 2 = 0.976 \rightarrow 0$
3	$x(2) = 0$	$y(2) = 5/8$	$(0.593)_{16}$ $= (0.10010)$	$(0.1010)_2 = 5/8$	$y(3) = 5/8$	
4	$x(2) = 0$	$y(2) = 5/8$	$(0.593)_{16}$ $= (0.10010)_2$	$(0.1010)_2 = 5/8$	$y(4) = 5/8$	
5	$x(2) = 0$	$y(2) = 5/8$	$(0.593)_{16}$ $= (0.10010)_2$	$(0.1010)_2 = 5/8$	$y(5) = 5/8$	

www.Vidyarthiplus.com
www.Vidyarthiplus.com

www.Vidyarthiplus.com

$$\delta(n) = \begin{cases} 1 & n=0 \\ 0 & \text{otherwise} \end{cases}$$

$$x(n) = \begin{cases} 3/4 & n=0 \\ 0 & \text{otherwise} \end{cases}$$

Deadband :

$\alpha = 0.95$

$$\leq \frac{1/2 \; 2^{-4}}{1-0.95}$$

$$\leq \frac{0.03125}{0.05}$$

$$\leq 0.625$$

$$\leq 5/8$$

www.Vidyarthiplus.com