Digital Audio Effect

by
Syed Siraj-Ul-Haque

Submitted in accordance with the requirements for the degree of
MSc(Eng) in Modern Digital and Radio Frequency Wireless Communications

University of Leeds
School of Electronic and Electrical Engineering

September, 2004

Abstract

The aim of this project is to design a "Digital audio effects" black box (as an acronym we use DAFX), which takes input audio signal, modify input signal characteristic by some control parameters to create desired effect and deliver output signal. The input and output audio signals are monitored by acoustical and visual representation of the signal. These DAFX s' can be realised by implementing suitable digital filters whose response can be controlled.

In short following are the goal of this project

- Classification and definition of audio effects.
- Design algorithm/filters which can create these effects on Audio signals.
- Study the change in response by changing control parameters.
- Provide graphical control panel for controlling filter parameters.
- Provide acoustical and visual representation of input and output audio signals.

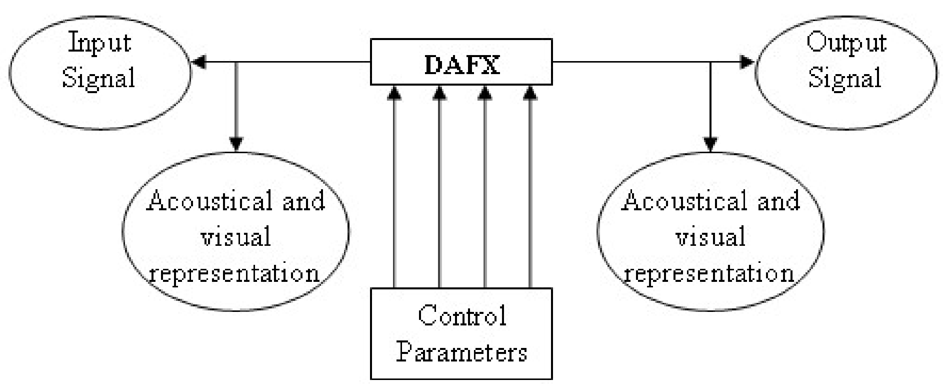


Figure Block diagram of DAFX

Contents

Figures

Tables

Flow Charts

Preface

Chapter I

Representation of Audio Signals

Chapter 1
Representation of Audio Signals

In this project six different type of acoustical and visual representation of audio signal is provided. Each representation is used for different purposes. (These representations could be seen in snap short portion of report)

- Audio Output
- Time domain representation
- Magnitude Spectrum (Frequency Spectrum)
- Phase Spectrum (Frequency Spectrum)
- Waterfall Representation
- Spectrogram

1.1. Audio Output

In order to hear audio signal and effect applied on it "playing option" is available to user.

1.2. Time Domain Representation

It's an "amplitude verses time" graph of the given signal.

1.3. Frequency Spectrum

Fourier transform can be used to find out the frequency domain representation of a time domain signal. The appropriate formulation for our discrete signal is the DTFT. For implementation and simulation purposes faster version of DTFT called fast Fourier transformation is used. The inverse Fourier transform converts a frequency domain representation into time domain.

Fourier's Transform is a valuable analytical tool because it allows breaking a complex signal down into a collection of simple signals. When applied to two complex signals, comparing the properties of the two resulting collections of simple signals is a useful way to analyse the similarities and differences between the two signals.

The amplitude of the signal as a function of frequency is referred to as amplitude spectrum and the phase of the signal as a function of frequency is referred to as phase spectrum of the spectrum of the signal. The amplitude and phase spectra together are called the frequency spectrum of the signal.

In MATLAB FFT function is used to compute the Fourier transform, the resulting vector will contain amplitude and phase information on positive and negative frequencies. Command used for amplitude spectrum is Y= abs(fft(y)) and for phase spectrum is Y= angle(fft(y)).

1.4. Waterfall Plot

Waterfall plot is a three dimensional representation of spectral changes over time. (Frequency spectra as a function of time). Water fall plot is made by series of short time Fourier transform taken at regular intervals of time. Early transform move down the display followed by later maps, something like the flow of a waterfall.

Water fall plot and SFFT is different from FFT because FFT only indicates the presence of various frequencies within the signal. It doesn't include information about when different frequencies occur or how they change over time. STFT and waterfall plot addresses this by computing the FFT for small overlapping windows of audio. This produces a slowly time-varying FFT.

1.5. Spectrogram

The spectrogram is a computer-generated display that produces a visual representation of the partials of an acoustical signal over time.

A spectrogram is a visual representation of the frequency content of a signal. A spectrogram shows how the quantity of energy in different frequency regions varies as a function of time. On a spectrogram, the signal is divided into many small time sections and each section is analyzed in terms of what frequency components are present in the section. This analysis is called spectral analysis because the spectrum of each section is calculated and the quantity of each frequency component (that is each sinusoid) is measured from the spectrum. The quantity of each component is then converted to a grey level in which (normally) low energy components are converted to a white colour, while high energy components are converted to a black colour. These colours are then plotted on a vertical strip corresponding to the time at which the original signal segment occurred. The height of the coloured element on this vertical strip represents the frequency of the component.

Thus a spectrogram is a 3-dimensional analysis of a signal, the horizontal dimension is time, the vertical dimension is frequency, and the grey-scale shows the amount of energy occurring in the signal at each time and frequency.

Time runs from left to right. The frequencies of the harmonics are indicated from bottom to top (thus, the higher the horizontal line on the image, the higher the frequency of that component).

Each horizontal line on the graphic represents a harmonic in the signal and their relative strengths can be read by changes in colour and brightness. Formants (concentrations of strong resonance in the sound) show up as the brightest bands.

Chapter 2

Effects Based on variation of Time Delay

Chapter 2
Effects Based on variation of Time Delay

These effects are based on the addition of time-delayed samples to the current output. The simulator designed for these effects is based on a structure proposed by Dattorro [Dat97] ("Dattorro structure" is used in this report to refer to this structure). In order to understand this understand working of this structure we have to under stand working of FIR, IIR & All-Pass comp filter plus fractional delay line.

2.1. Basic Delay Structure

2.1.1 FIR Comb Filter

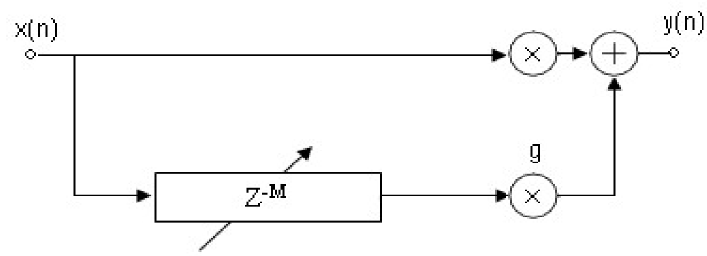


Figure -2.1 FIR Comb Filter

- This network produces single echo in input signal. It adds M sample delayed version of input signal to original input signal.
- The effect is audible only when the processed signal is combined (added) to input signal (which acts as a reference) and M is sufficiently larger enough.
- This effect has two tuning parameters
  1. Amount of delay τ seconds
  2. Relative amplitude of delayed signal to that of the reference signal 'g'.
- Difference equation and transfer function of FIR Comb Filter

$$y(n) = x(n) + gx(n-M) \quad \text{With} \quad M = \tau / f_s \qquad 2.1$$

$$H(z) = 1 + gz^{-M} \qquad 2.2$$

- Effect of gain ‘g’ on spectrum
    - If g is positive, the filter amplifies all frequencies that are multiples of $1/\tau$ and attenuates all frequency that lie in between.
    - If g is negative, the filter attenuates all frequencies that are multiples of $1/\tau$ and amplifies all frequency that lie in between.
    - The gain varies between 1+g and 1-g.
- Effect of FIR comb filter in time domain and frequency domain
    - For Large values of τ, we can hear echo but can not notice spectral effect of comb filter
    - For Small values of τ, our ears can not segregate time events but can notice the spectral effects

2.1.2 IIR Comb Filter

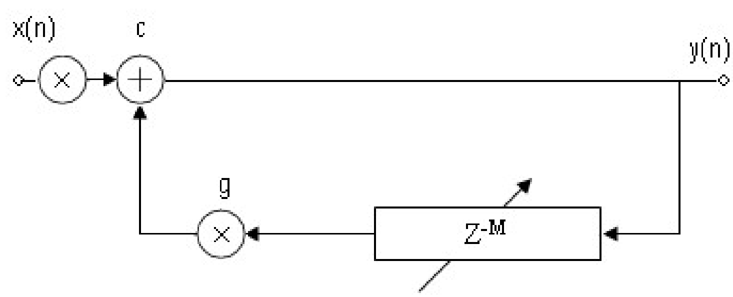


Figure 2.2 IIR Comb Filter

- IIR comb filter produces endless series of echo in input signal.
- Each time the signal goes through the delay line that is fed back to the input, it is attenuated by g.
- It is sometime necessary to scale the input by ‘c’ in order to compensate for the high amplification produced by the structure.
- Difference equation and transfer function of IIR Comb Filter

$$y(n) = cx(n) + gy(n-M) \quad \text{With} \quad M = \tau / f_s \qquad 2.3$$

$$H(z) = c/(1 + gz^{-M}) \qquad 2.4$$

- After each time delay τ a copy of the input will come out with amplitude of g^P (P is number of cycles).
- IIR Comb Filter is stable as long as |g| <= 1. Other wise signal will grow endlessly.
- The gain varies between 1/(1-g) and 1/(1+g)
- The main difference between FIR and IIR comb is that that the gain grows very high and that frequency peaks get narrower as |g| comes close to 1.

2.1.3 Universal Comb Filter

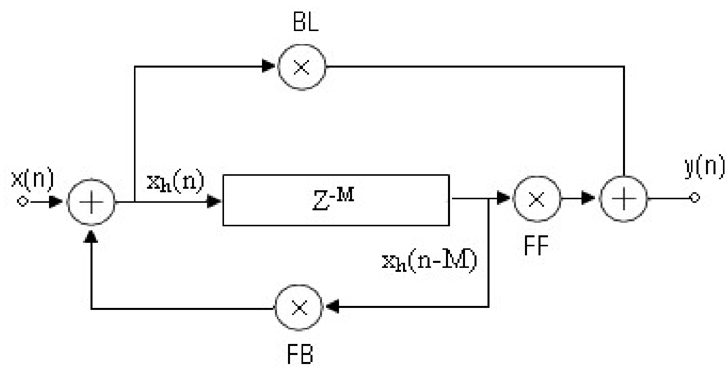


Figure 2.3 All-pass Comb Filter

- This filter is cascaded combination of FIR & IIR comb filter which share same delay line.
- Difference equation of the universal Comb filter

$$x_h(n) = x(n) + FBx_h(n-M) \tag{2.5}$$

$$y(n) = BLx(n) + FFx_h(n-M) \tag{2.6}$$

- These equations show a digital filter in direct form II. Difference equation of direct form I of same filter is given below. (See diagram of direct form I & II on next page).

$$y(n) = BLx(n) + FFx(n-M) + FBy(n-M) \tag{2.7}$$

- Transfer function of the universal Comb filter

$$H(z) = (BL + FFz^{-M})/(1 + FBz^{-M}) \tag{2.8}$$

- Use of different parameter sets leads to the different applications shown below.

	BL	***FB***	***FF***
FIR Comb Filter	X	0	X
IIR Comb Filter	1	X	0
All-Pass	a	-a	1
Delay	0	0	1

Table 2.1 Parameters for universal comb filter

	Delay	***BL***	***FB***	***FF***
Slap back	50ms	1	0	X
Echo	>50ms	1	0<X<1	0
Reverb		Matrix	Matrix	Matrix

Table 2.2 Effects with generalized comb filter

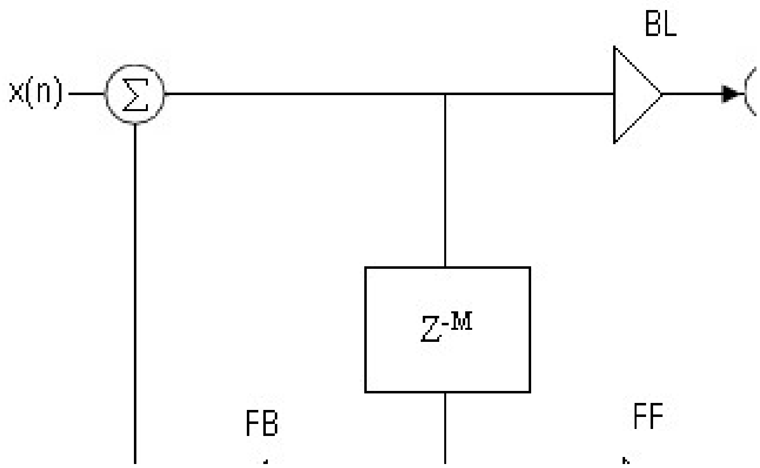


Figure 2.4 Direct form II of universal comb filter

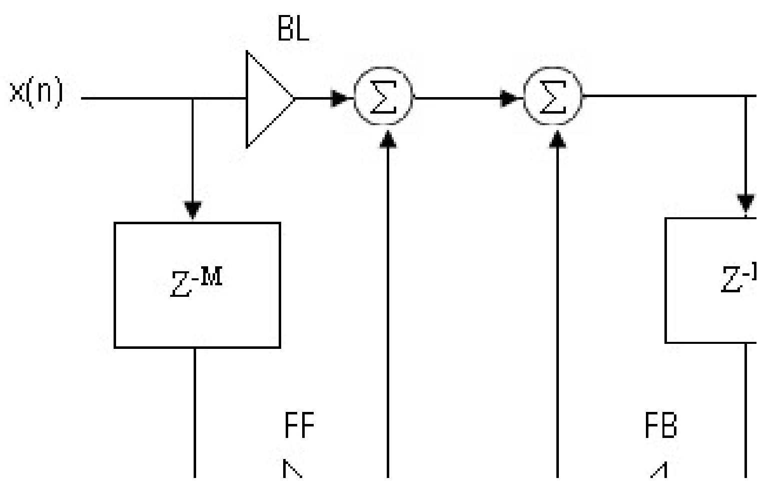


Figure 2.5 Direct form I of universal comb filter

2.1.4 Fractional Length Delay Line

- This delay line introduces non integer value of sample delay.
- A delay of the input signal by M samples plus a fraction of the normalized sampling interval with 0<farc<1 is given by. $y(n) = x(n - [M + frac])$
- An interpolation algorithm is used to compute the output sample y(n), which lies in between the two samples at time instants M and M+1.
- Several interpolation algorithms have been proposed for audio applications. The choice of the algorithm depends on the specific application. Some of them are mentioned below.
    1. Linear interpolation

       $$y(n) = x(n - [M + 1])frac + x(n - M)(1 - frac) \qquad 2.9$$

    2. Allpass interpolation

       $$y(n) = x(n - [M + 1])frac + x(n - M)(1 - frac) - y(n - 1)(1 - frac) \qquad 2.10$$

    3. sinc interpolation
    4. fractionally addressed delay lines
    5. spline interpolation
- Interpolation techniques only work for low frequency signal.

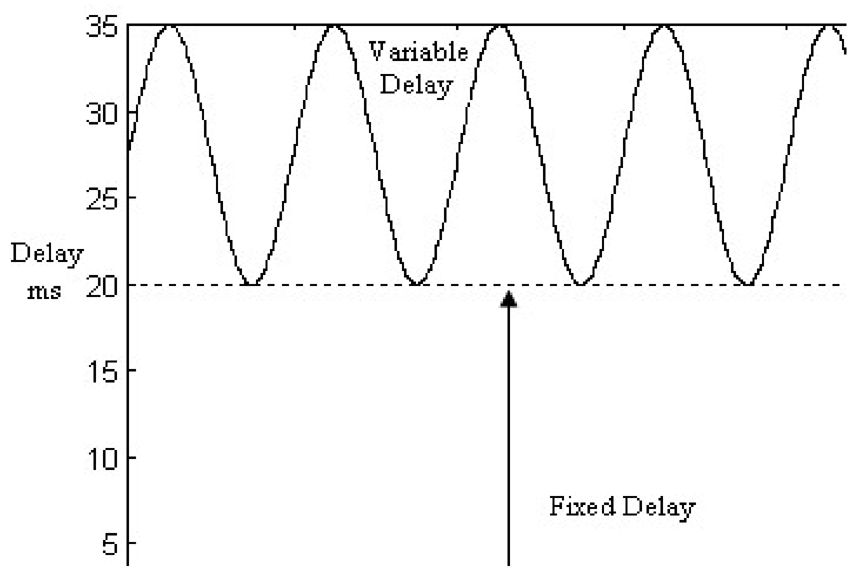


Figure 2.6 Sinusoidal varying delays

2.1.5 - Dattorro Standard Effect Structure

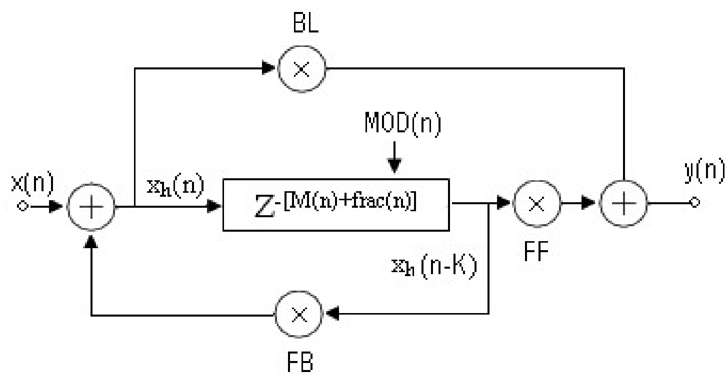


Figure 2.7 Dattorro standard effect structure

- This structure was proposed by Dattorro [Dat97].
- It is based on the all-pass filter modification towards a general all-pass comb, where the fixed delay line is replaced by a variable length delay line.
- Dattorro proposed keeping the feedback tap of the delay line fixed, that means the input signal to the delay line x_h (n) is delayed by a fixed integer delay K and with x_h (n-k) is weighted and fed back.
- The delay K is the centre tab delay of the variable length delay line for the feed forward path.
- The control signal MOD(n) for changing the length of the delay line can either be a low frequency sinusoid or low pass noise.
- Typical setting of the parameters are given below

	BL	***FF***	***FB***	***Delay***	***Depth***	***MOD***
Vibrato	0	1	0	0 ms	0-3 ms	0.1-5 Hz
Flanger	0.7	0.7	0.7	0 ms	0-2 ms	0.1-1 Hz
Chorus	0.7	1	-0.7	1-30 ms	1-30 ms	Low pass Noise
Doubling	0.7	0.7	0	10-100 ms	1- 100 ms	Low pass Noise

Table 2.3 Parameter for Dattorro standard effect structure

2.1.6 How to remove metallic sound of Comb filter

- Comb filter sounds metallic because they tend to amplify greatly the high frequency component.
- In order to make comb filter sound more natural, the attenuation factor ‘g’ has be frequency dependent.
- This frequency dependence can be realized by using a first order low pass filter in the feedback loop.

2.2 Delay Effects

2.2.1 Slap back Effect

If the delay is in the range of 10 to 25ms, we will hear a quick repetition named slap back or doubling. Slap back effect with a low decay value added to a voice track can make the voice sound "metallic" or robot-like. This was a popular way of creating the robotic-voice in movies in days gone by.

2.2.2 Echo

Echo is produced by adding a time-delayed signal with delay greater then 50ms to the output. A single echo can be produced by FIR filter. Multiple echoes are achieved by IIR filter. It is observed that there is a threshold value of the delay, above which the echo can be clearly heard and below which the echo is absent. Echo greatly improves the sound of a distorted lead guitar solo, because it improves sustain and gives an overall smoother sound.

2.2.3 - Chorus

If several copies of the input signal are delayed in the range 10 to 25ms with small and random variations in the delay time, we will hear chorus effect. Note that the delay is varied continuously between a minimum delay and maximum delay at a certain rate. The chorus effect is so named because it makes the recording of a vocal track sound like it was sung by two or more people singing in chorus.

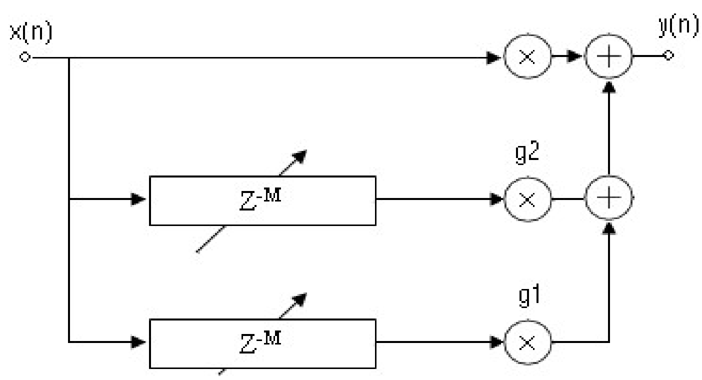
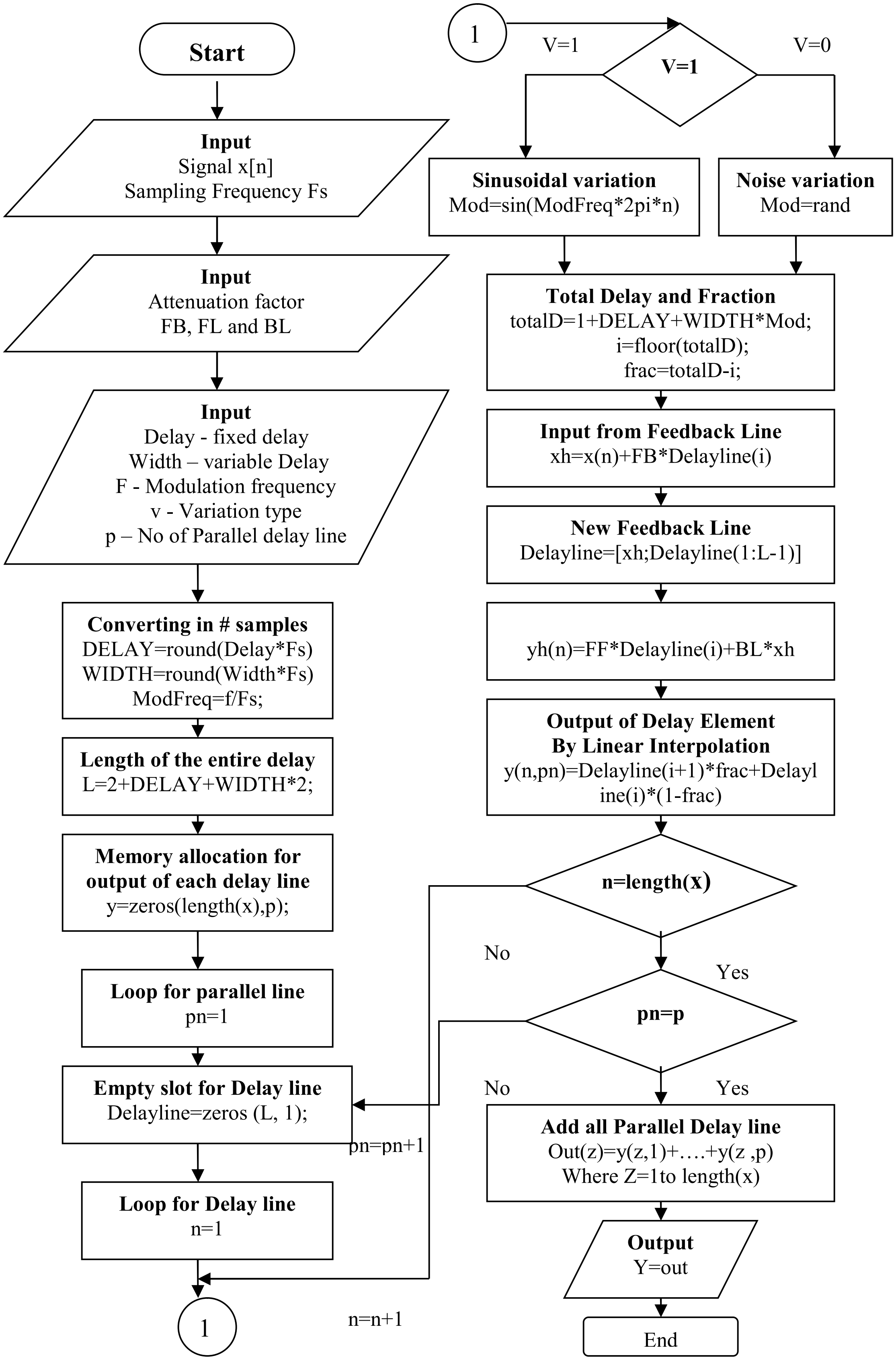


Figure 2.8 Structure for Chorus

2.2.4 - Flanging Effect

Flanging is a special case of the chorus effect: it is created in the same way that chorus is created. Typically, the delay of the echo for a flanger is varied between 0ms and 15ms at a rate of 0.5Hz.

In days gone by, flanging used to be created by sound engineers who put their finger onto the tape reel's flange, thus slowing it down. Two identical recordings are played back simultaneously, and one is slowed down to give the flanging effect.



Flow Chart -1: Working of Dattorro Standard Effect Structure

2.2.5 - Reverberation

Reverb is used to simulate the acoustical effect of rooms and enclosed buildings. Reverb is basically a combined effect of multiple sound reflections within the room.

The reverberation characteristics of a room are effected by several factors like shape size material of which the room is constructed and the materials present in room. These materials are especially important since they determined how much sound is absorbed and how much is reflected.

In a room, for instance, sound is reflected off the walls, the ceiling and the floor. The sound heard at any given time is the sum of the sound from the source, as well as the reflected sound. An impulse (such a hand clap) will decay exponentially. The reverberation time is defined as the time taken for an impulse to decrease by 60dB of its original magnitude. Echo is distinct delayed version of the signal whereas a reverb has a very short delay time that makes it difficult to distinguish each reflection. Concert halls and rooms have to be designed such that the reverberation time is adequate for the type of sound that will be produced. Reverberation time which is adequate for orchestral music can be reached only if the volume per seat (in a theatre) is about $5m^3$ per person, which is expensive to achieve. Thus, concert halls tend to have too short a reverberation time. Reverberation time of a hall can be lengthened by using a digital reverberator which adds reverb to the sound, and then re-radiates it in the original room by loudspeaker arrays.

From the first Schroeder work (early 1960) about artificial reverberations based on discrete time signal processing, many implementations have been presented in different publications. Each reverberation algorithm has specific characteristics like natural sounding, absence of tonal coloration, high echo density control of the reverberation time etc.

We will consider a typical impulse response, associated with the reverberation effect, where we can distinguish the early echoes and the late reverberation. We will use Moorer’s implementation of Schroeder reverberation algorithm to simulate this effect.

The block A of Moorer reverb simulates early reflections by mean of direct form FIR filter, usually this FIR filter is implemented as a tapped delay line, i.e. a delay line with multiple reading points that are weighted and summed together to provide a single output.

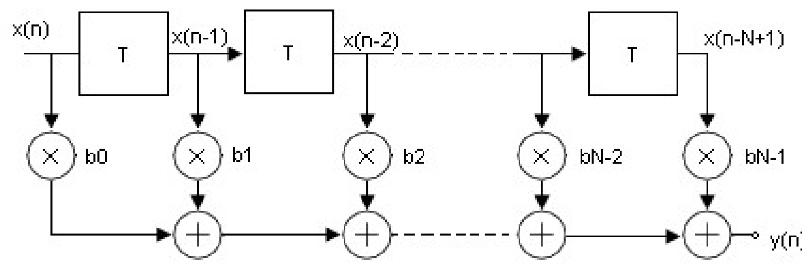


Figure 2.9 ‘Block A’: Tapped Delay Line FIR filter

The resulting signal is forwarded to the block B, which is the parallel of a direct path on one branch and a delayed attenuation diffused reverberator on the other branch. In Moorer’s preferred implementation the reverberator of block B is best implemented as a parallel group of six comb filter. Each comb filter has a first order low pass filter in their feedback loop. These low pass filter are used to simulate effects of air absorption and lossy reflections [Moo79] it also removes metallic character of sound.

Moorer suggested setting the all pass filter delay length to 6ms and all pass filter coefficient to 0.7. The feedback attenuation coefficient should be set in a way which results in smooth decay. If the desired decay time (usually defined for an attenuation level of 60db) is T_d and m_i is delay length, the gain of each comb filter has to be set to [Moo79].

$$g_i = 10^{-3\frac{T_d F_S}{m_i}} \qquad 2.11$$

The output of reverberator is delayed in such a way that the last of the early echoes coming out of the block reaches the output before the first of non null samples coming out of the diffused reverberator.

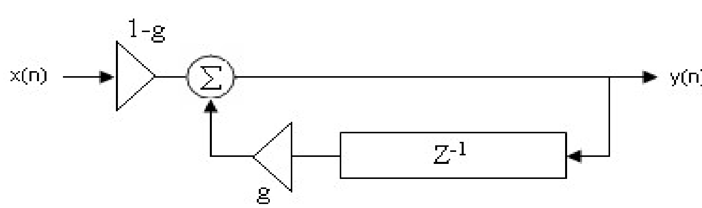


Figure 2.10 Low Pass filter block diagram

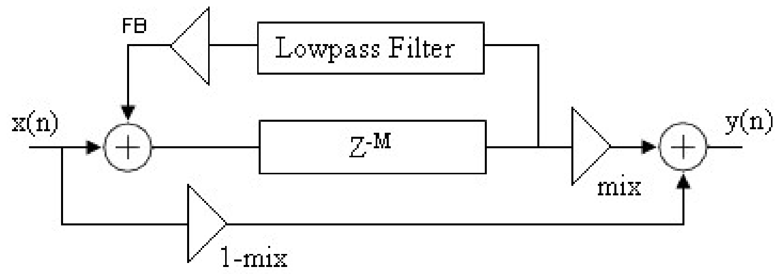


Figure 2.11 Comb filter with low pass filter in feedback loop

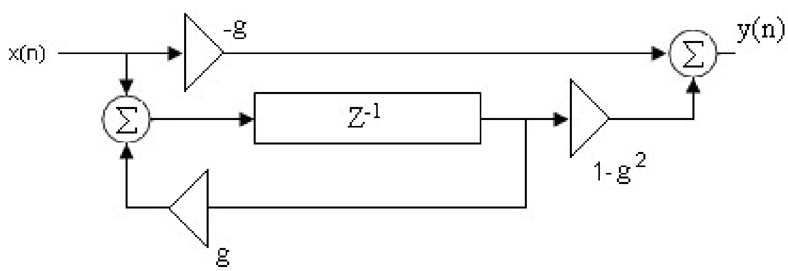


Figure 2.12 All-pass Reverberation Filter

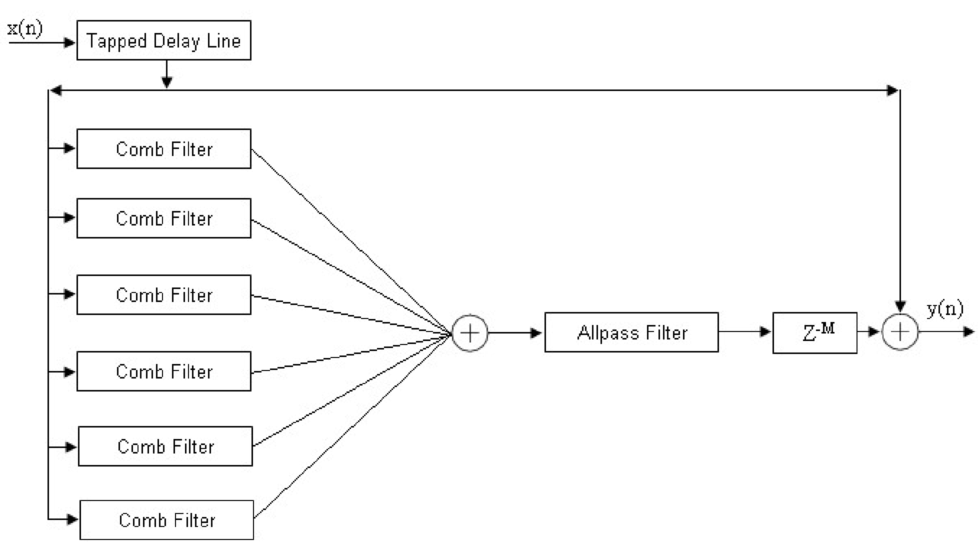
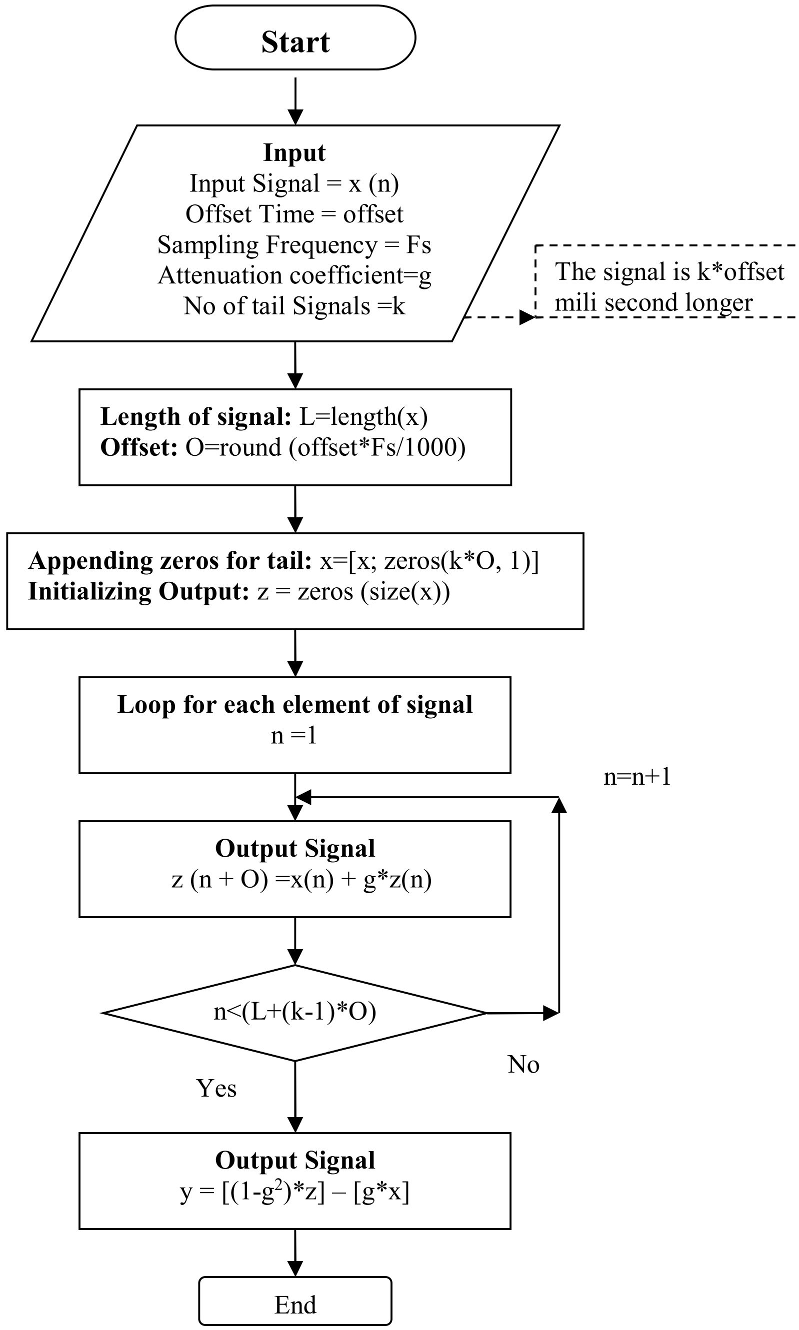
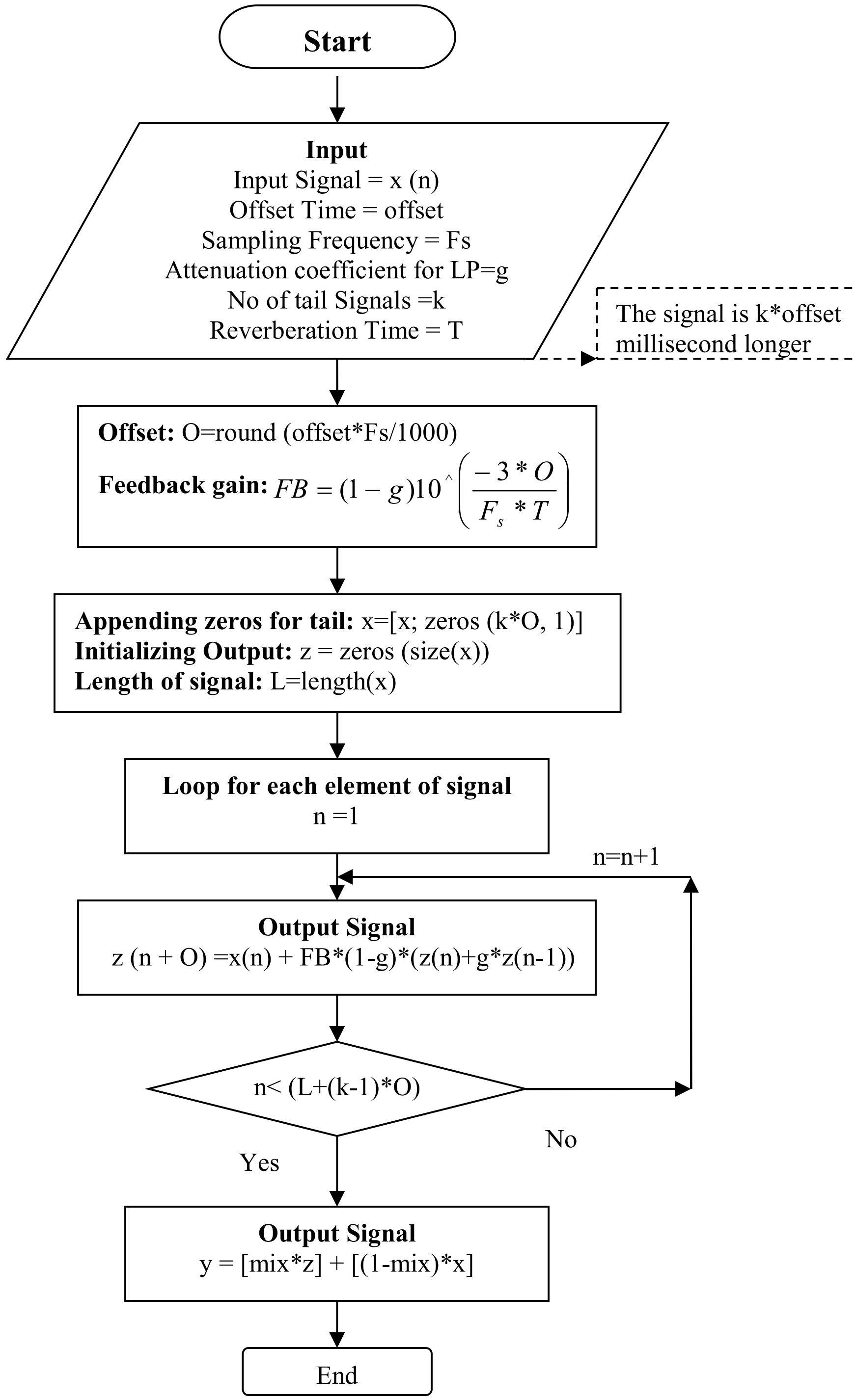


Figure 2.13 ‘Block B’: Moorer’s Reverberation [Moo79]

Figure 2.14 Impulse response of a Room. Earlier Reflection is simulated by Block A and later reflection is simulated by Block B



Flow Chart 2.2 Working of All-pass Reverberation Filter Simulator



Flow Chart 2.3 Working of "Comb Filter with Low Pass Filter" Simulator

Chapter 3

Effects based on variation of Amplitude

Chapter 3
Effects based on variation of Amplitude

These are effects based on variation of amplitude

- Tremolo
- Ring Modulation
- Panning
- Distance Rendering
- Fading
- Compression/Expansion (Dynamic controller)
- Limiter (Dynamic controller)
- Noise Gate (Dynamic controller)

3.1. Tremolo

Cyclical variation of volume by a low frequency oscillator of some sort; parameters are waveform of the LFO, LFO frequency, and depth of modulation; note that while the terms tremolo and vibrato are often used interchangeably, tremolo is actually variation in loudness, vibrato is variation in pitch or frequency.

Auto tremolo: Tremolo where the modulation frequency is varied by some feature of the input signal, generally amplitude.

3.2. Ring Modulation

A ring modulator is a simple device that can be used to create unusual sounds from an instruments output. It effectively takes two signals (each with some frequency), and produces a signal containing the sum and differences of those frequencies. These frequencies will typically be non-harmonic, so the ring modulator can create some very dissonant sounds. For this reason, ring modulation is not a widely used effect. The audio signal x(n) is multiplied by a sinusoid m(n) with carrier frequency f_c.

$$y(n) = x(n) \,.\, m(n) \qquad 3.1$$

Output y(n) is made up of two copies of the input spectrum LSB & MSB. LSB is reversed in frequency and both sidebands are centred on f_c. Depending on the width of the spectrum of x(n) and on the carrier frequency, the side bands can be partly mirrored around the origin of the frequency axis. If the carrier signal comprises several spectral components the same effect happens with each component.

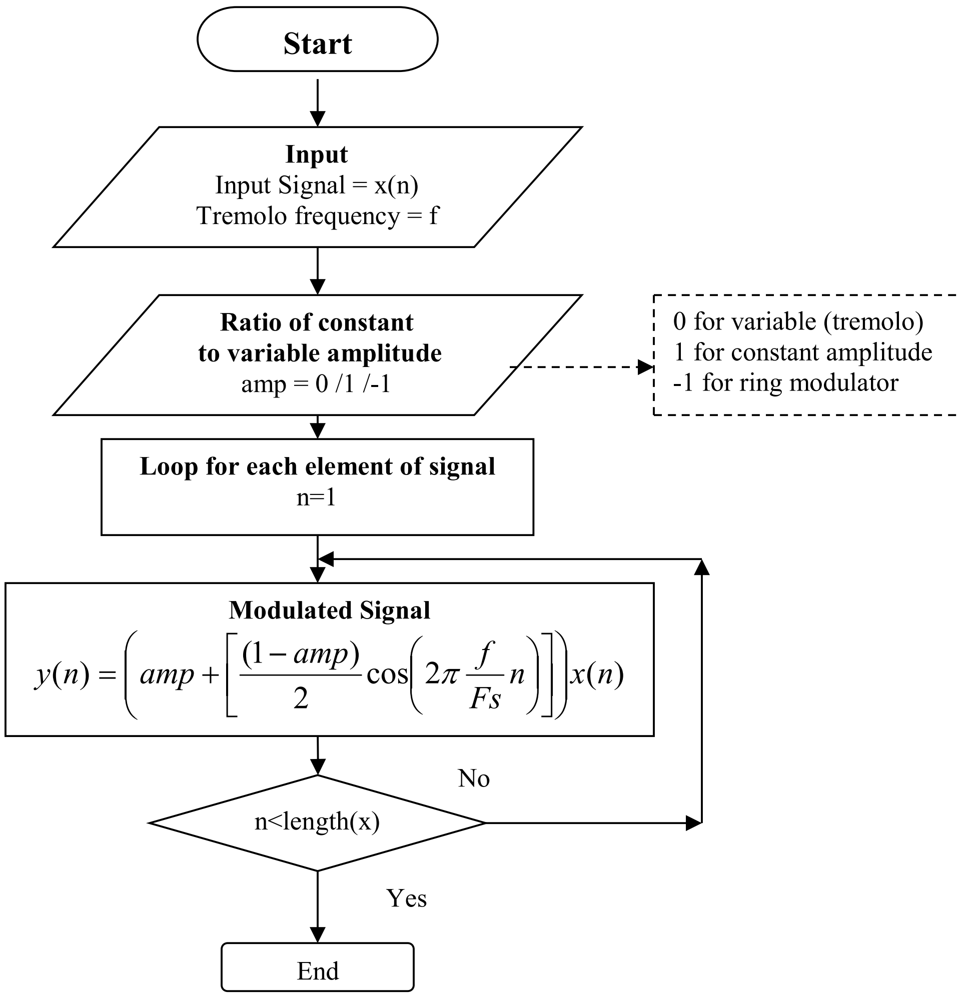
Although the audio result of a ring modulation is fairly easy to comprehend for elementary signals. It gets very complicated with signals having numerous partials. The carrier itself is not audio in this kind of modulation.

When carrier and modulator are sine waves of frequency f_c and f_x one hears the sum and the difference frequencies $f_c + f_x$ & $f_c - f_x$.

When the input signal is periods with fundamental frequency f_0, a sinusoid carrier of frequency f_c produces a spectrum with amplitude line at the frequencies $[kf_0 +/- f_c]$.

3.3. Fading

Fading in/out is performed by simply modulating given signal by growing/ decaying function. These growing or decaying functions could be linear or exponential. In this project linear function is used.



Flow Chat 3.1 Working of Tremolo and Ring modulation

2.2.5 - Reverberation

Reverb is used to simulate the acoustical effect of rooms and enclosed buildings. Reverb is basically a combined effect of multiple sound reflections within the room.

The reverberation characteristics of a room are effected by several factors like shape size material of which the room is constructed and the materials present in room. These materials are especially important since they determined how much sound is absorbed and how much is reflected.

In a room, for instance, sound is reflected off the walls, the ceiling and the floor. The sound heard at any given time is the sum of the sound from the source, as well as the reflected sound. An impulse (such a hand clap) will decay exponentially. The reverberation time is defined as the time taken for an impulse to decrease by 60dB of its original magnitude. Echo is distinct delayed version of the signal whereas a reverb has a very short delay time that makes it difficult to distinguish each reflection. Concert halls and rooms have to be designed such that the reverberation time is adequate for the type of sound that will be produced. Reverberation time which is adequate for orchestral music can be reached only if the volume per seat (in a theatre) is about $5m^3$ per person, which is expensive to achieve. Thus, concert halls tend to have too short a reverberation time. Reverberation time of a hall can be lengthened by using a digital reverberator which adds reverb to the sound, and then re-radiates it in the original room by loudspeaker arrays.

From the first Schroeder work (early 1960) about artificial reverberations based on discrete time signal processing, many implementations have been presented in different publications. Each reverberation algorithm has specific characteristics like natural sounding, absence of tonal coloration, high echo density control of the reverberation time etc.

We will consider a typical impulse response, associated with the reverberation effect, where we can distinguish the early echoes and the late reverberation. We will use Moorer's implementation of Schroeder reverberation algorithm to simulate this effect.

The block A of Moorer reverb simulates early reflections by mean of direct form FIR filter, usually this FIR filter is implemented as a tapped delay line, i.e. a delay line with multiple reading points that are weighted and summed together to provide a single output.

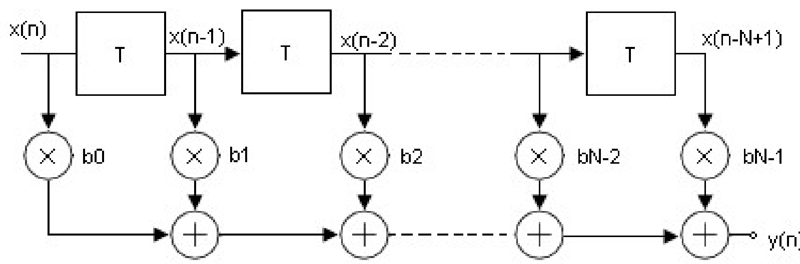


Figure 2.9 'Block A': Tapped Delay Line FIR filter

The resulting signal is forwarded to the block B, which is the parallel of a direct path on one branch and a delayed attenuation diffused reverberator on the other branch. In Moorer's preferred implementation the reverberator of block B is best implemented as a parallel group of six comb filter. Each comb filter has a first order low pass filter in their feedback loop. These low pass filter are used to simulate effects of air absorption and lossy reflections [Moo79] it also removes metallic character of sound.

Moorer suggested setting the all pass filter delay length to 6ms and all pass filter coefficient to 0.7. The feedback attenuation coefficient should be set in a way which results in smooth decay. If the desired decay time (usually defined for an attenuation level of 60db) is T_d and m_i is delay length, the gain of each comb filter has to be set to [Moo79].

$$g_i = 10^{-3\frac{T_d F_S}{m_i}} \qquad 2.11$$

The output of reverberator is delayed in such a way that the last of the early echoes coming out of the block reaches the output before the first of non null samples coming out of the diffused reverberator.

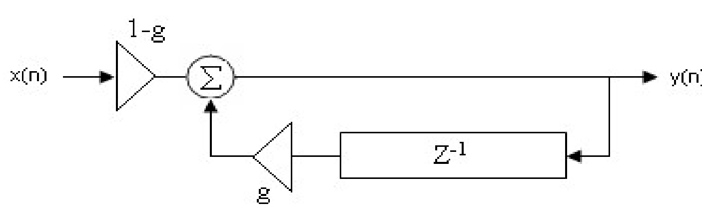


Figure 2.10 Low Pass filter block diagram

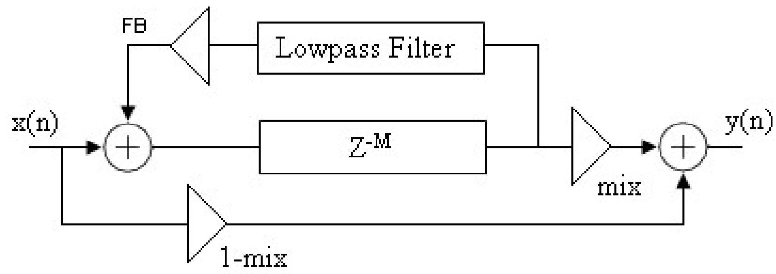


Figure 2.11 Comb filter with low pass filter in feedback loop

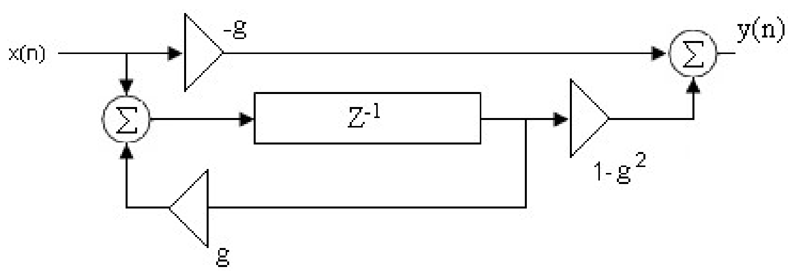


Figure 2.12 All-pass Reverberation Filter

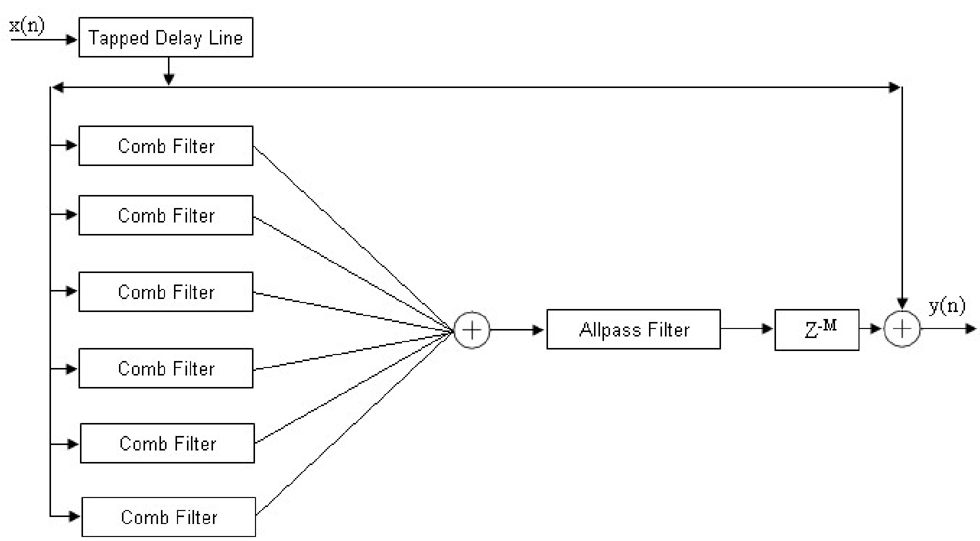


Figure 2.13 'Block B': Moorer's Reverberation [Moo79]

Figure 2.14 Impulse response of a Room. Earlier Reflection is simulated by Block A and later reflection is simulated by Block B

Chapter 3

Effects based on variation of Amplitude

Chapter 3
Effects based on variation of Amplitude

These are effects based on variation of amplitude

- Tremolo
- Ring Modulation
- Panning
- Distance Rendering
- Fading
- Compression/Expansion (Dynamic controller)
- Limiter (Dynamic controller)
- Noise Gate (Dynamic controller)

3.1. Tremolo

Cyclical variation of volume by a low frequency oscillator of some sort; parameters are waveform of the LFO, LFO frequency, and depth of modulation; note that while the terms tremolo and vibrato are often used interchangeably, tremolo is actually variation in loudness, vibrato is variation in pitch or frequency.

Auto tremolo: Tremolo where the modulation frequency is varied by some feature of the input signal, generally amplitude.

3.2. Ring Modulation

A ring modulator is a simple device that can be used to create unusual sounds from an instruments output. It effectively takes two signals (each with some frequency), and produces a signal containing the sum and differences of those frequencies. These frequencies will typically be non-harmonic, so the ring modulator can create some very dissonant sounds. For this reason, ring modulation is not a widely used effect. The audio signal x(n) is multiplied by a sinusoid m(n) with carrier frequency f_c.

$$y(n) = x(n) \,.\, m(n) \qquad 3.1$$

Output y(n) is made up of two copies of the input spectrum LSB & MSB. LSB is reversed in frequency and both sidebands are centred on f_c. Depending on the width of the spectrum of x(n) and on the carrier frequency, the side bands can be partly mirrored around the origin of the frequency axis. If the carrier signal comprises several spectral components the same effect happens with each component.

Although the audio result of a ring modulation is fairly easy to comprehend for elementary signals. It gets very complicated with signals having numerous partials. The carrier itself is not audio in this kind of modulation.

When carrier and modulator are sine waves of frequency f_c and f_x one hears the sum and the difference frequencies $f_c + f_x$ & $f_c - f_x$.

When the input signal is periods with fundamental frequency f_0, a sinusoid carrier of frequency f_c produces a spectrum with amplitude line at the frequencies $[kf_0 +/- f_c]$.

3.3. Fading

Fading in/out is performed by simply modulating given signal by growing/ decaying function. These growing or decaying functions could be linear or exponential. In this project linear function is used.

3.4 Panorama Effect (Stereo Panning)

Using a “multi-channel sound reproduction system” we can change the apparent position of a source just by “feeding the channels with the same signal” and “adjusting the relative amplitude of the channels”. By doing so these signals arrives at slightly different levels and times and we get virtual sound image at a different place from the real source locations.

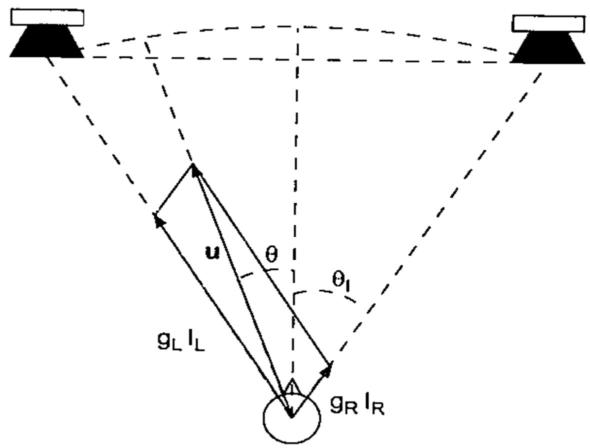


Figure 3.1 Stereo panning

θ=angle of virtual source

θ_L=angle formed by each loudspeaker with the frontal direction

g_L & g_R =gain applied to left & right stereo channels

In a standard loudspeaker setup it is assumed that listener stands in central position and form an angle $2\theta_L$ with the two loudspeakers. Two gains g_L & g_R are applied in order to set apparent azimuth at the desired value θ.

A unit magnitude “two channel signal corresponding to the central apparent source position (θ = 0) can be represented by the column vector.

$$U = \begin{bmatrix} \sqrt{2}/2 \\ \sqrt{2}/2 \end{bmatrix} \qquad 3.2$$

So gains to be applied to the two channels in order to steer the sound source to the desired azimuth are obtained by the matrix vector multiplication.

$$A_\theta U = \begin{bmatrix} g_L \\ g_R \end{bmatrix} \qquad \text{Where } A_\theta\text{=Rotation matrix} \qquad 3.2$$

Amplitude panning by means of a rotation matrix preserves the loudness of the virtual sound source while moving its apparent azimuth.

It θ_L=45 the rotation matrix takes the form shown below, so that θ= +/- θ_L only one of the two channels is non zero. It is easily verified that the rotation by matrix given below corresponds to applying the tangent law to the configuration with θ_L=45°.

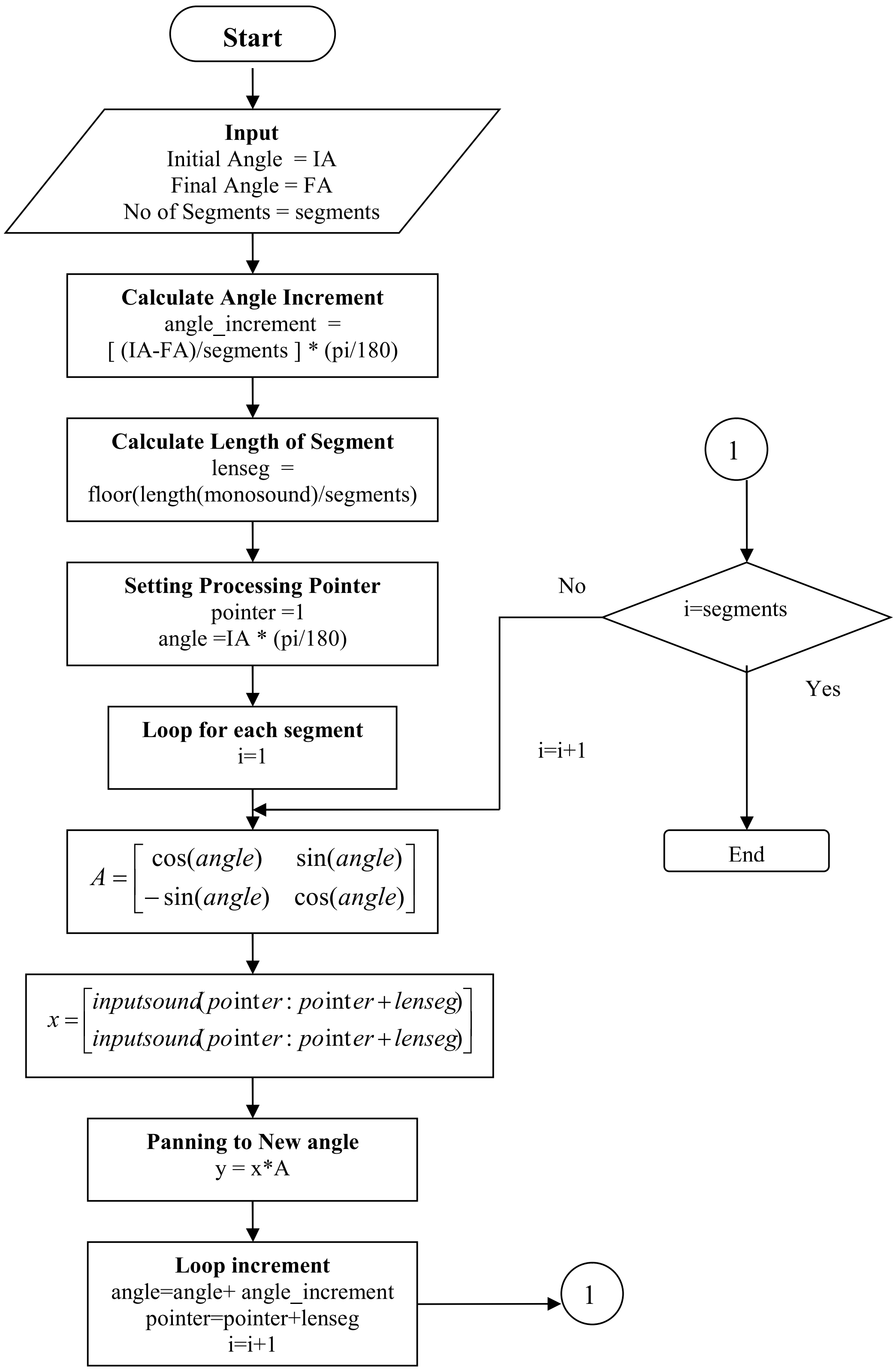
$$A_\theta = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \qquad 3.3$$

Tangent Law (results from vector formulation of amplitude panning):

$$\tan\theta = \frac{g_L - g_R}{g_L - g_R}\tan\theta_L \qquad 3.4$$

In practice the steering angle θ does not necessarily correspond to the perceived localization azimuth. The perceived location is influenced by the frequency content of the sound. Some theories of directional psychoacoustic have been developed in the past in order to drive the rotation matrix with appropriate coefficients .Accurate implementation use frequency dependent rotation matrix, at least discriminating between low (less than about 500Hz) and high (between 500Hz and 3500 Hz) frequency components.

In implementation input sound is segmented in a number “segment” of blocks and each block is rotated of a matrix by vector multiplication. Flow chart on next page represents implements the amplitude panning between an initial and a final angle.



Flow Chart 3.2: Panning Effect algorithm

3.5 Distance Rendering

In digital audio effects the control of apparent distance can be effectively introduced even in monophonic audio systems. In fact the impression of distance of a sound source is largely controllable by insertion of artificial wall reflections or reverberant room response.

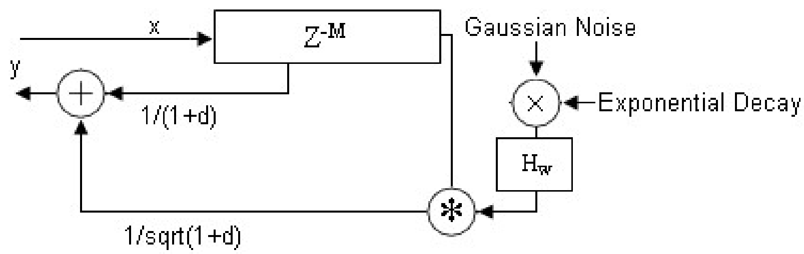


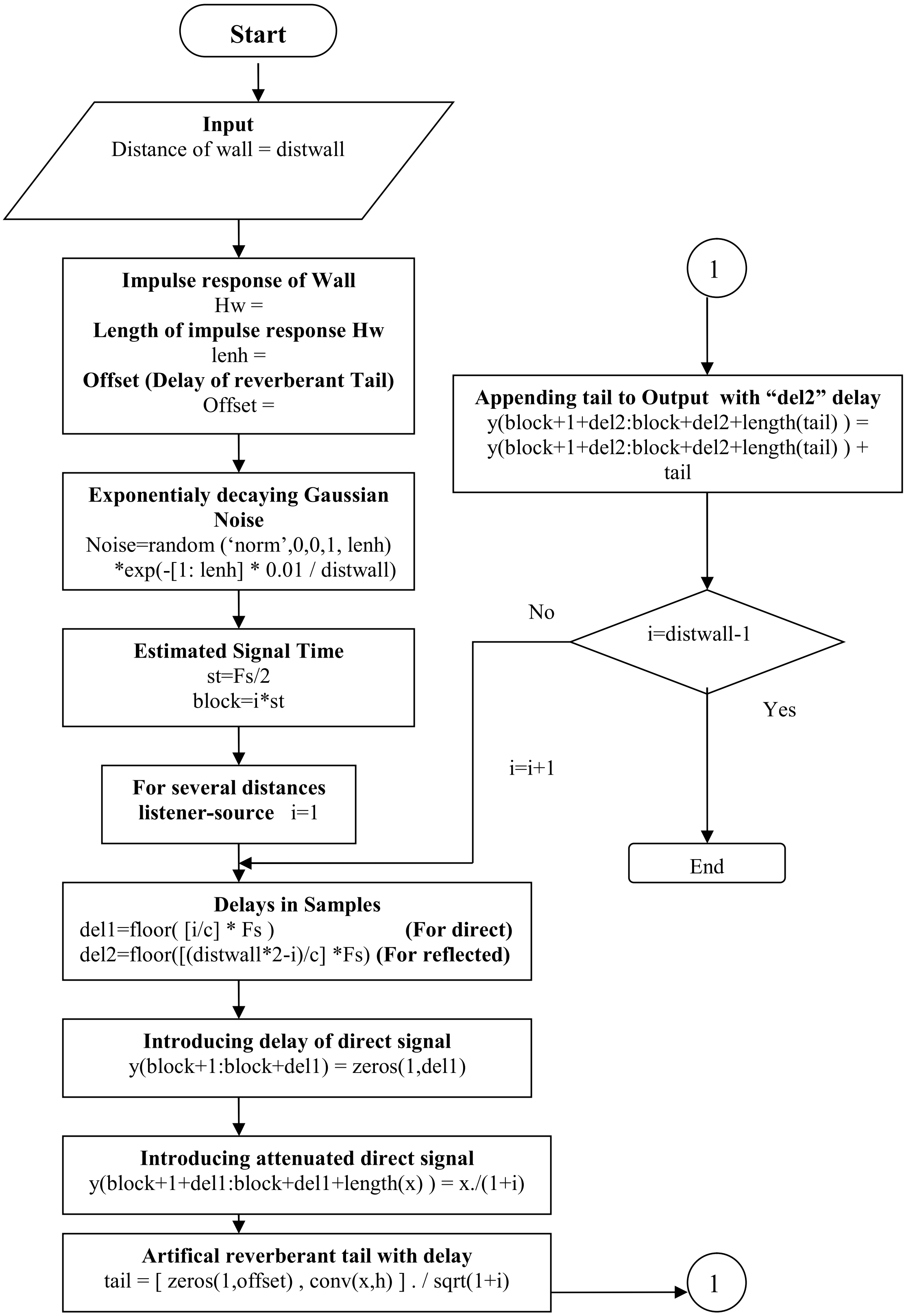
Figure 3.2 DSP Design for Distance rendering

There are no reliable cues for distance in open space but in enclosure the ratio of reverberant to direct acoustic energy has proven to be robust distance cue.

It is assumed that a small space the amplitude of the reverberant signal changes little with distance and that in a large space it is roughly proportional to $1/\sqrt{Distance}$ [Cho71]. The direct sound attenuates as 1/distance if spherical waves are propagated.

- A single reflection from a wall can be enough to provide some distance cues in many cases.
- A single delay line with two taps is enough to reproduce this basic effect.
- If the virtual sound source is close enough to the listening point the first tap can be taken directly from the source thus reducing the signal processing circuitry to single non-recursive comb filter.
- To be physical consistent the direct sound and its reflection should be attenuated as much as the distance they travel.
- The wall reflection should also introduce some additional attenuation and filtering in the reflected sound (represented by H_w).

- The distance attenuation coefficient is set in such a way that they become one when the distance goes to zero just to avoid the divergence to infinity that would come from the physical laws of a point source.
- From this simple situation it is easy to see how the direct sound attenuates faster than the reflected sound as long as the source approaches the wall. This idea can be generalized to closed environments adding a full reverberant tail to the direct sound.
- An artificial sound yet realistic reverberant tail can be obtained just by taking an exponentially decayed Gaussian noise and convolving it with the direct sound.
- The reverberant tail should be added to the direct sound after some delay (proportional to the size of the room) and should be attenuated with distance in a lesser extent than the direct sound.



Flow Chart 3.3: Distance Rendering algorithm

3.6 Dynamics processing

3.6.1. Concept behind Dynamic Processing

Dynamic processing is based on following

- An amplitude detection scheme called “envelop follower”.
- An algorithm to drive a gain factor from the result of the “envelope follower”
- A multiplier to weight the input signal.

The envelope follower calculates the mean of the absolute value over predefined time interval. Certain thresholds are defined for a change of the output to input behaviour.

The lower path of “dynamic controller” consists of the envelope detector and the following processing to drive the gain factor is usually called the **side chain path**. The output of the side chain path is used as gain factor for the delayed input signal (coming from direct path). Normally the gain factor is derived from the input signal but the side chain can also be connected to another signal for controlling the gain factor of the input signal.

Dynamic controller consists of

- Direct path for delaying input signal
- Side Chain path

Side chain path perform a level measurement and a subsequently gain factor calculation which is then used as gain factor for the delayed input signal. The level measurement is followed by a static function & a part of for attack and release time adjustment. The calculation of the time variation gain factor g(n) is usually performed with a logarithm level representation because the human sensitivity of loudness follows a logarithmic relation.

The delay of samples in direct path allows for the time delay of the side chain processing which is mainly made up of the level measurement and the “attach and release” time adjustments.

The calculation of control parameter in the logarithms domain F in dB can be performed by simple line equations [Zol97, RZ95] given by

$$F_{LT} = \text{-LS(X - LT)} + \text{CS(CT - LT)} \qquad 3.5$$

$$F_{CT} = \text{-CS(X - CT)} \qquad 3.6$$

$$F_{LIN} = 0 \qquad 3.7$$

$$F_{ET} = \text{-ES(X - ET)} \qquad 3.8$$

$$F_{NG} = \text{-NS(X - NT)} + \text{ES(ET - NT)} \qquad 3.9$$

Where

LT=Limiting threshold

CT=Compressor threshold

ET Expander threshold

NT Noise Gate threshold

LS,CS,ES and NS are Slope factors (definition of Slope factor is given below)

Slope Factor $S = \frac{1}{1-R}$ 3.10

Compression factor $R = \frac{\Delta L_1}{\Delta L_o} = \frac{1}{1-S}$ 3.11

ΔL_I =Input Level change

ΔL_o =Output Level change

	Compression factor R	Slope Factor S
Limiter	R=∞	S=1
Compressor	1<R<∞	0<S<1
Linear Part	R=1	S=0
Expander	0<R<1	- ∞<S<0
Noise Gate	R=0	S=-∞

Table 3.1 Typical values of Slope factor and Compression Factor

The dynamic behaviour of a DR controller is influenced by the

- Level measurement approach
  1. **Peak measurement with AT & RT (Attack and Release Time)**

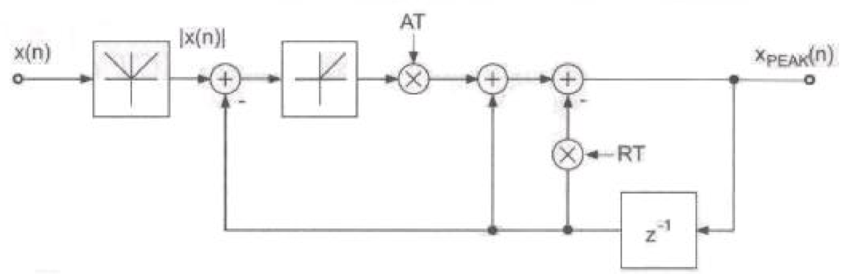


Figure 3.3 Peak Measurement

  2. **RMS measurement with TAV (Averaging time)**

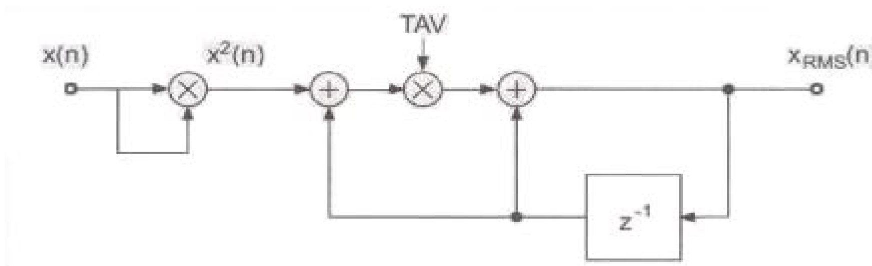


Figure 3.4 RMS measurement

- Adjustment of special attack/release times which can be achieved by the "filter with hysterisis" shown below

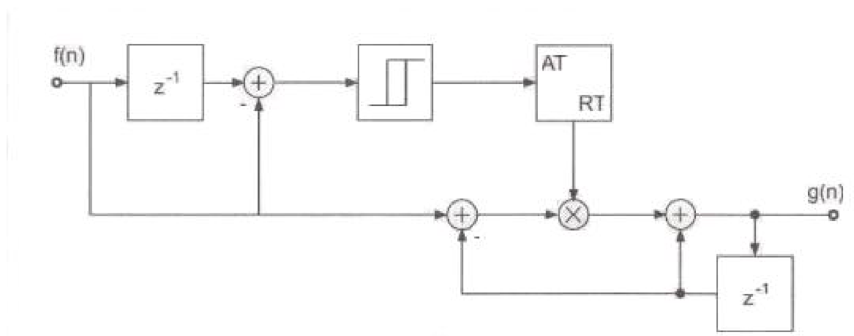


Figure 3.5 Dynamic Filter

The calculation of the attack time parameter is carried out by following equations [McN84, Zol97]

$$AT = 1 - e^{-2.2T/t_{AT}} \qquad 3.12$$

$$RT = 1 - e^{-2.2T/t_{RT}} \qquad 3.13$$

$$TAV = 1 - e^{-2.2T/t_{TAV}} \qquad 3.14$$

Where t_{AT} , t_{RT} & t_{TAV} are the time parameter and T is the sampling period. The output factor f(n) of the static function along with attack time and release time are used as the input signal to the dynamic filter. The output signal g(n) of dynamic filter is the gain factor for weighting the delayed input signal x(n-D).

3.6.2. Hysterisis

It is a phenomenon in which the response of a physical system to an external influence depends not only on the present magnitude of that influence but also on the previous history of the system.

If a gate is set with a fast attack time and a fast release time, any signal level which hovers around the threshold can end up causing a problem known as 'chattering', where the gate opens and closes rapidly several times in succession.

To avoid chattering hysterisis is used. Essentially, a hysteresis control raises the threshold for opening the gate and lowers that for closing it, such that they then differ by a few dBs. This means that, whatever the threshold level which opens the gate, the signal must fall a few dBs before the gate will be allowed to close again. As long as the release and hold times is set properly, hysteresis can help make gates behave much more smoothly and predictably. Following are the parameters used for hysterisis.

Upper threshold: Threshold value for deactivating the gate.

Lower threshold: Threshold value for activating the gate.

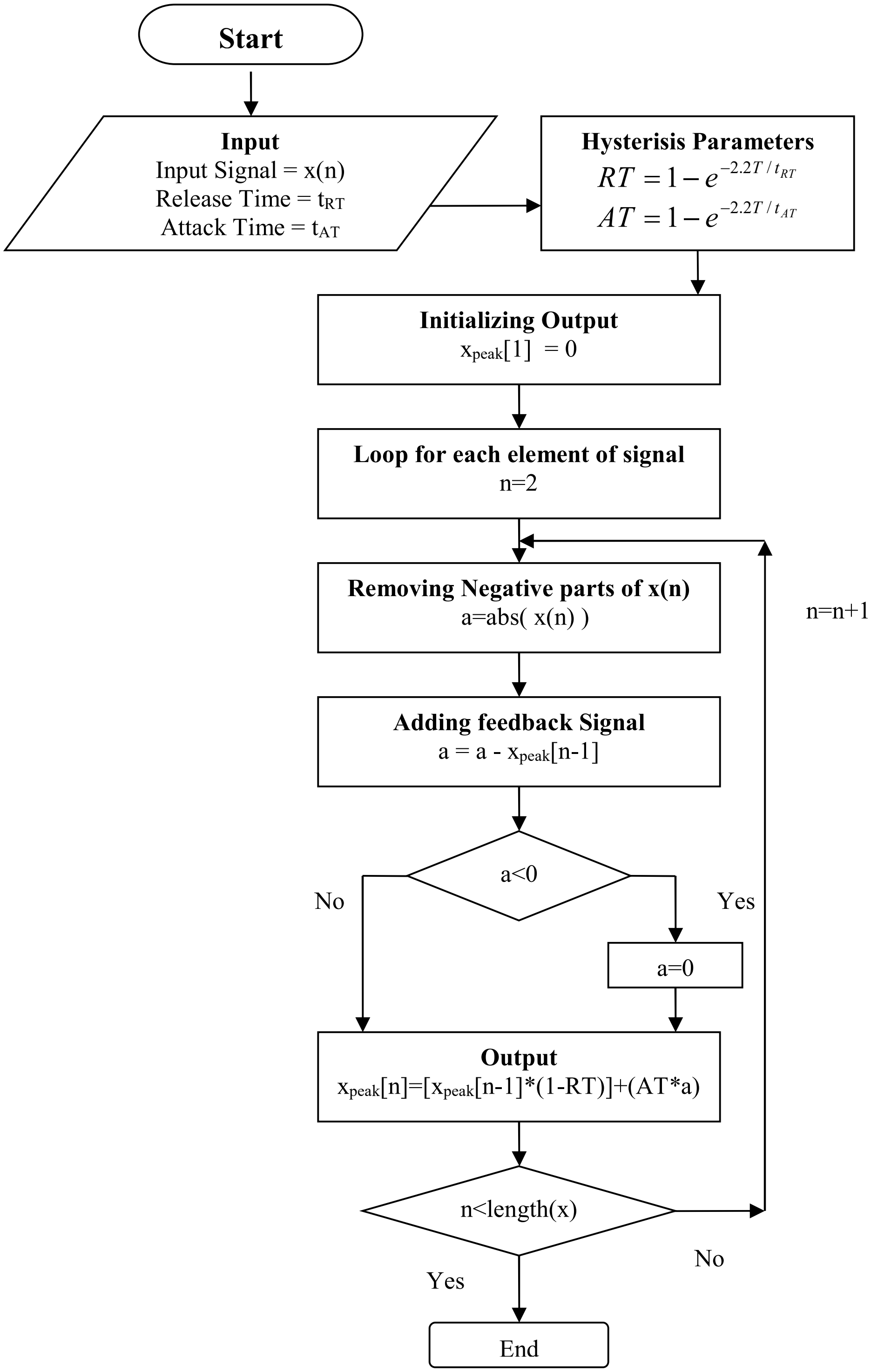
Hold time: Time for which a signal should remain below the threshold (lower) before gate is activated (signal is blocked). Hold Time control can be useful to stop the gate "chattering" open and closed when the source signal is constantly changing amplitude.

Attack Time: Time during which a signal greater than threshold (upper) is faded in from gated state. Time before the output level is the same as input level after deactivating the gate.

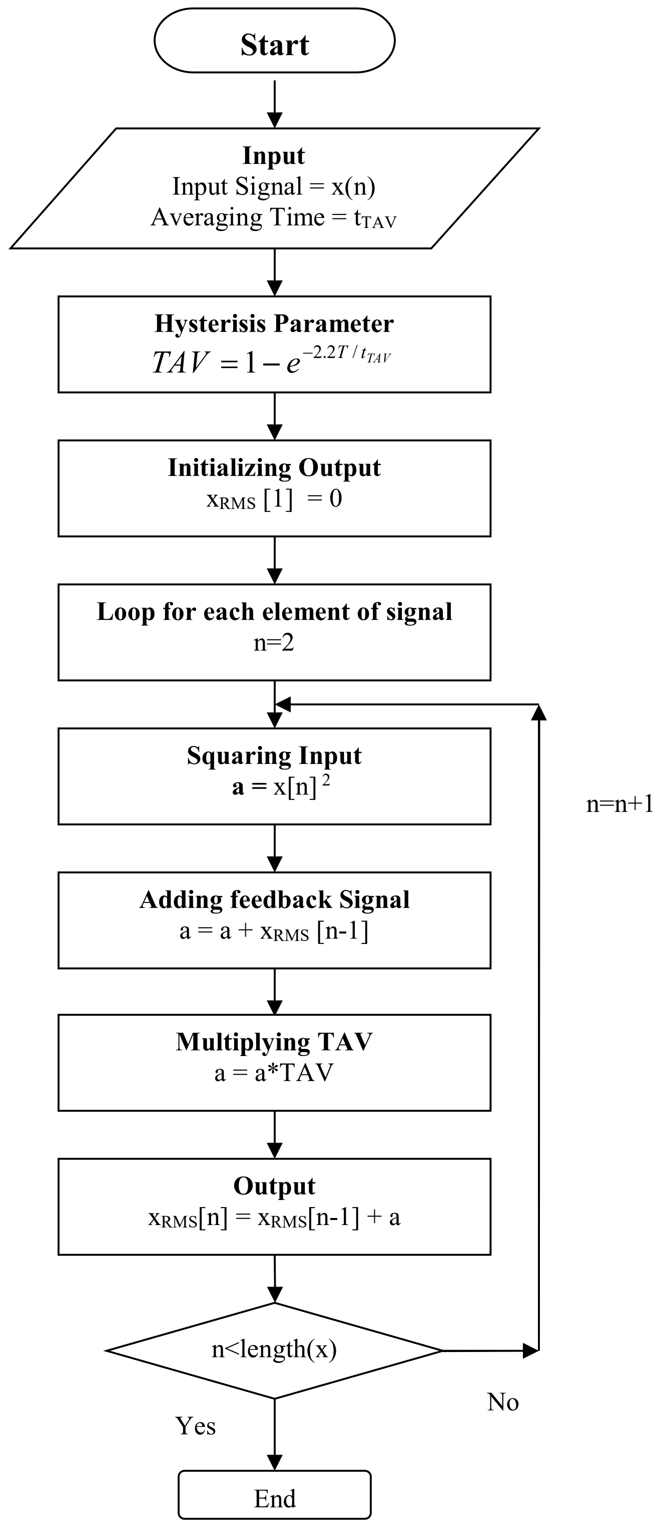
It is used by the user to set the time it takes for the gate to fully open after the signal exceeds the threshold level. Low frequency sounds are likely to be distorted if the gate opens too quickly. This is because the sudden change in gain could be applied to signal part of the way through the frequency cycle, resulting in an audible click. Conversely, if the Attack Time is too long, fast attack sounds, like a snare, will have less impact.

Release Time: Time during which a signal less than threshold (lower) is faded into gated state. Time before signal reaches zero.

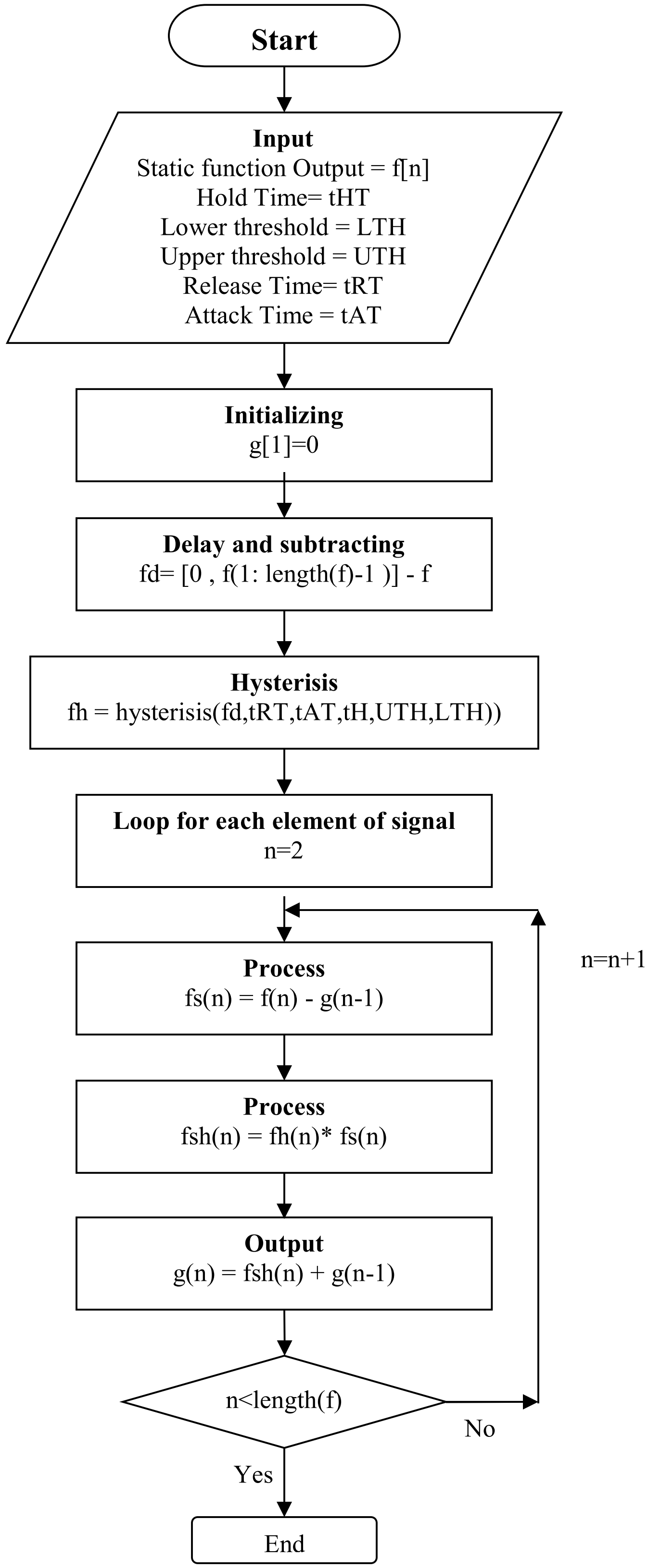
This allows the user to set the time over which the gate will close. This stops the sound or noise from suddenly stopping, instead the sound or noise gradually fades.



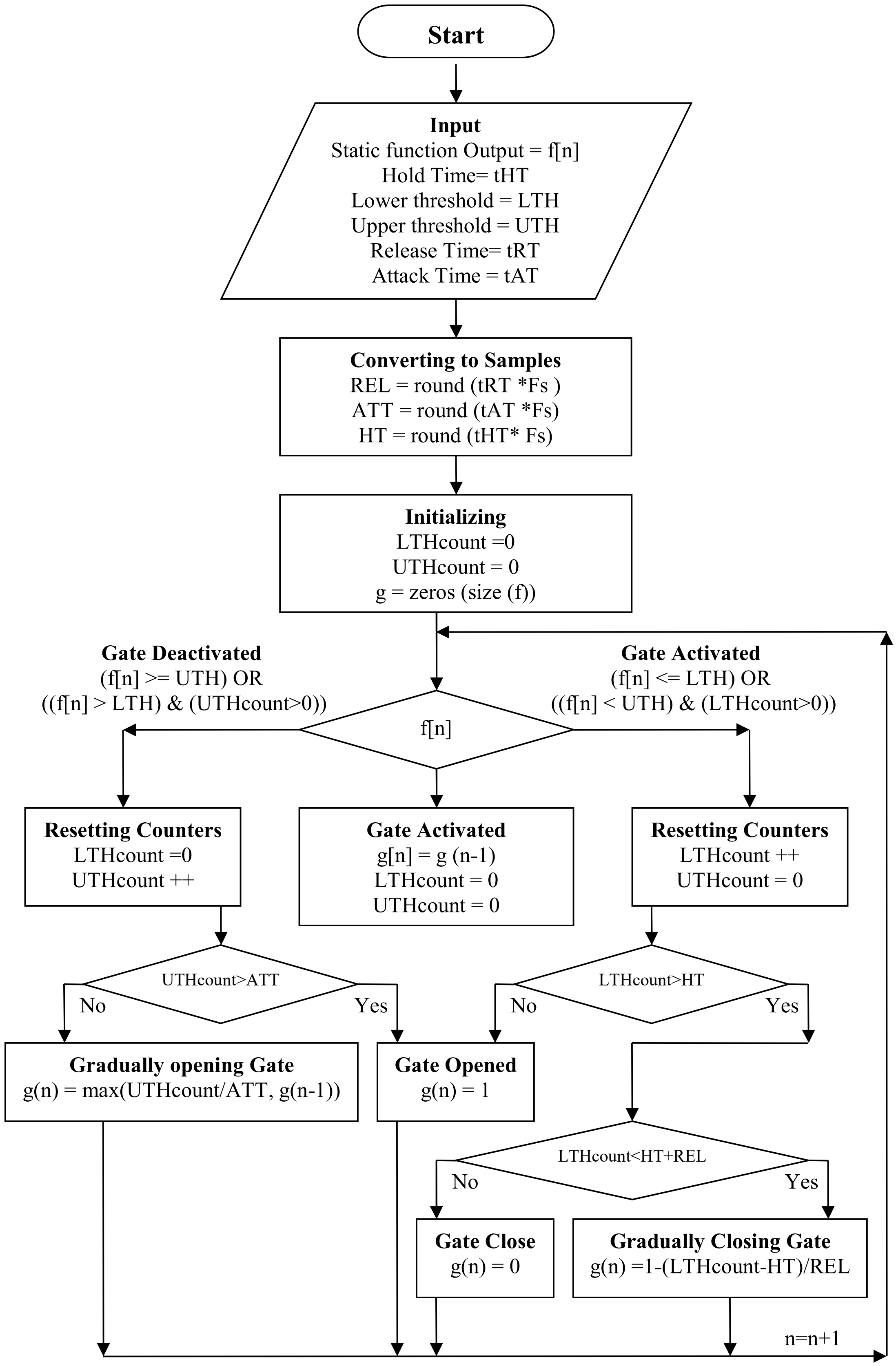
Flow Chart 3.4 Peak measurement with AT & RT



Flow Chart 3.5 RMS measurement with TAV



Flow Chart 3.6 Working of Dynamic Filter



Flow Chart 3.7 Working of Hysterisis

3.6.3. Limiter

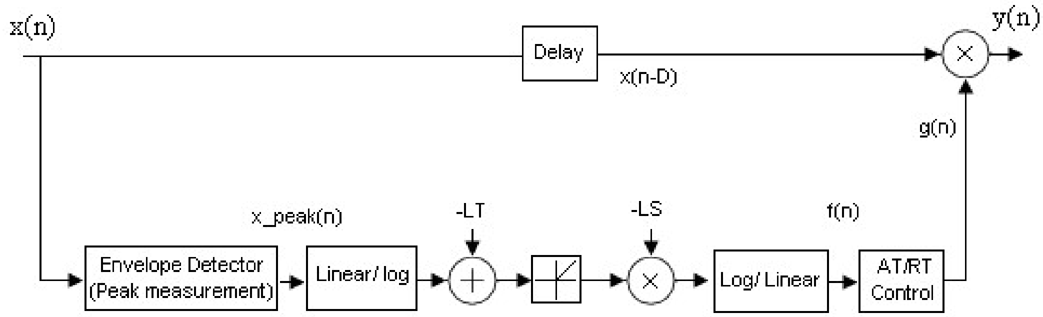


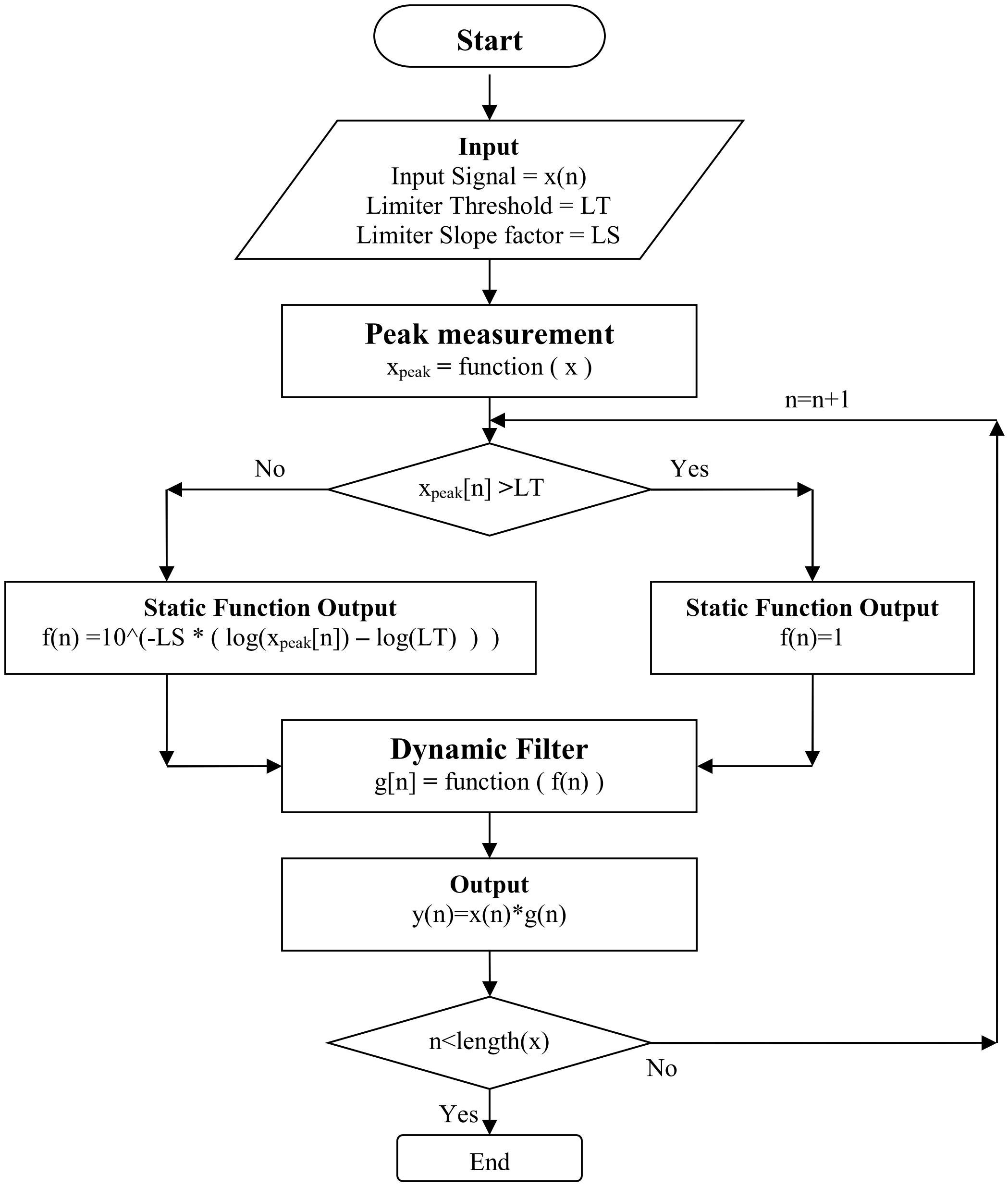
Figure 3.6 Structure of Limiter

The purpose of the limiter is to provide control over the high peaks in signal and to change the dynamics of the signals as little as possible.

Limiting Like compression, but operates on signals over some threshold only. Well suited to keep an input from going over some level, but un-processed below that level, as in getting signals on tape without overloading the tape.

A limiter makes use of peak level measurement and should reach very quickly to extension of the limiter threshold. The fast attack and release of a limiter allow the volume reduction as soon as the signal crosses the threshold. By lowering the peak the over all signal can be boosted.

Used for limiting signal instrument signals. Limiting is also often performed on the final mix of a multi channel application.



Flow Cart 3.8 Working of a Limiter

3.6.4. Compressor and Expander

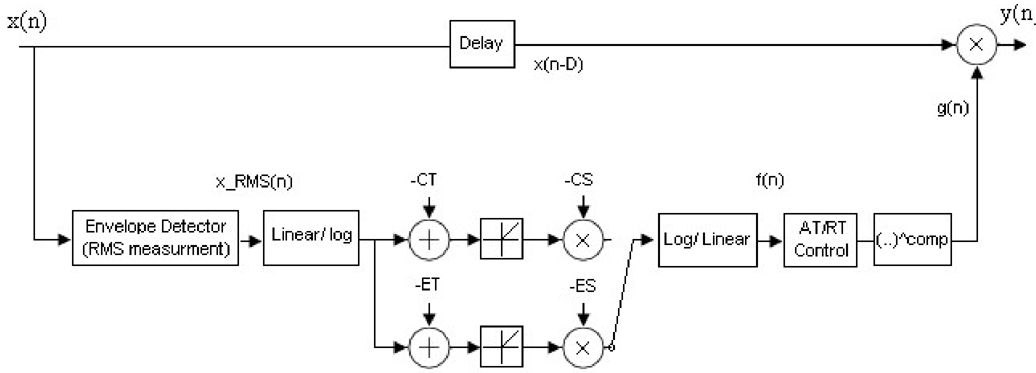
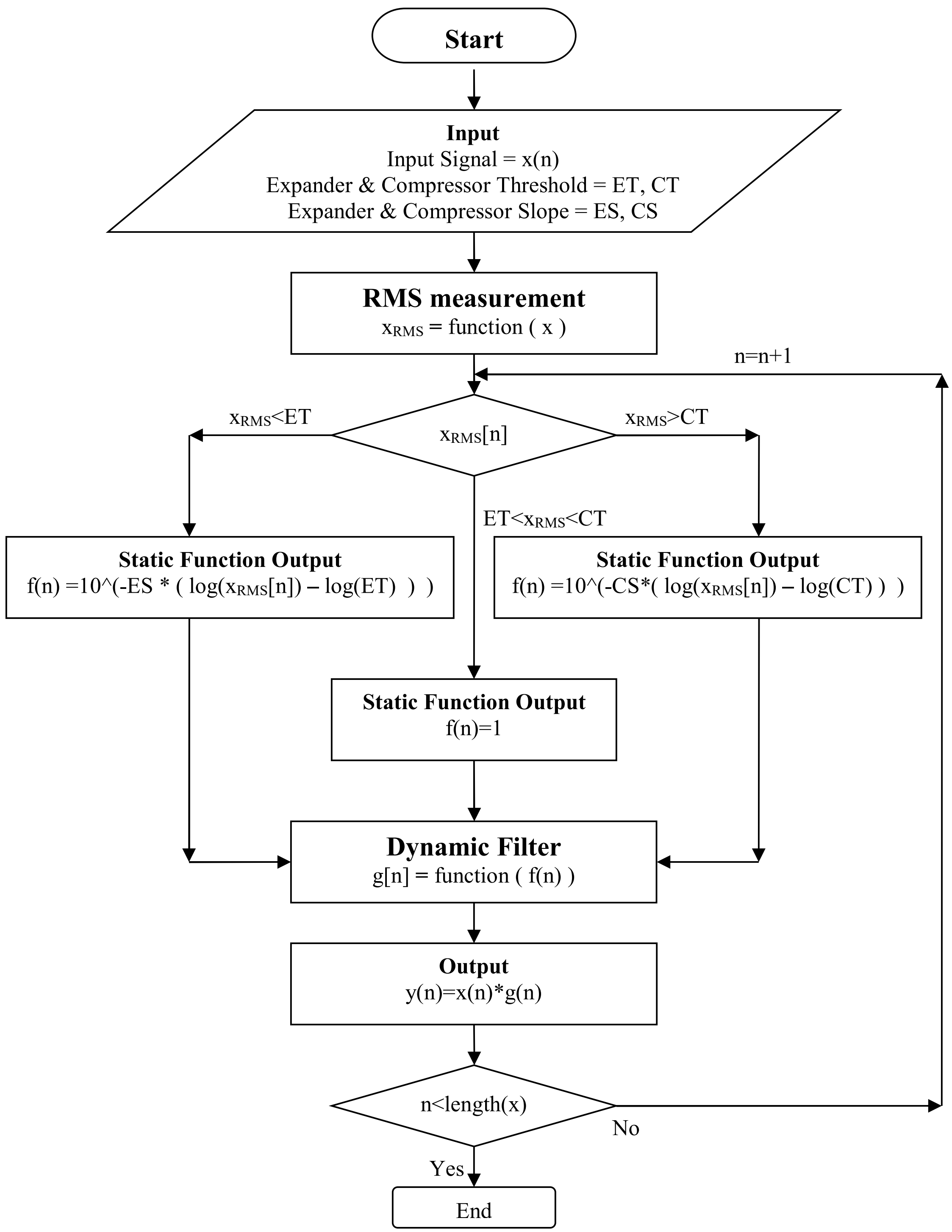


Figure 3.7 Structure of Compressor and Expander

The gain factor calculation is based on an RMS measurement and some computation in the logarithmic domain (described earlier).

Compressors are used for reducing the dynamics of the input signal. Low levels of a signal are not modified but high levels or are reduced according to the static curve. The result is that the difference between the loud and quiet parts is lessened and thus the overall signal level can be boosted and thus the signal is made louder.

Expanders operate on low level signals and increase the dynamics of these low level signals. This leads to a lively sound characteristic.



Flow Chart 3.9 Working of a Compressor / Expander

3.6.5. Noise Gate

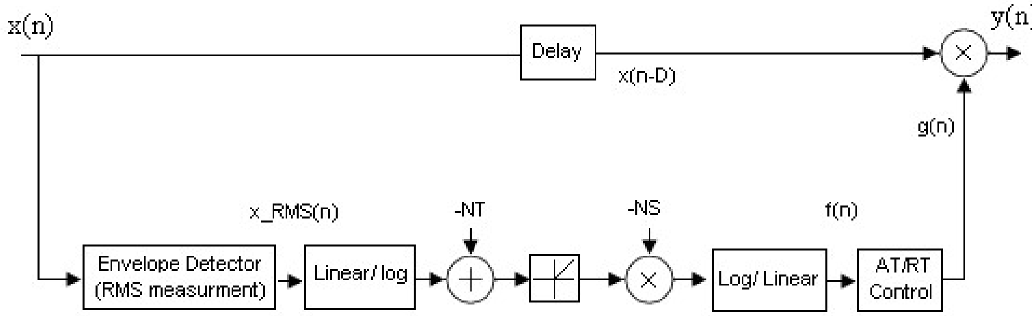


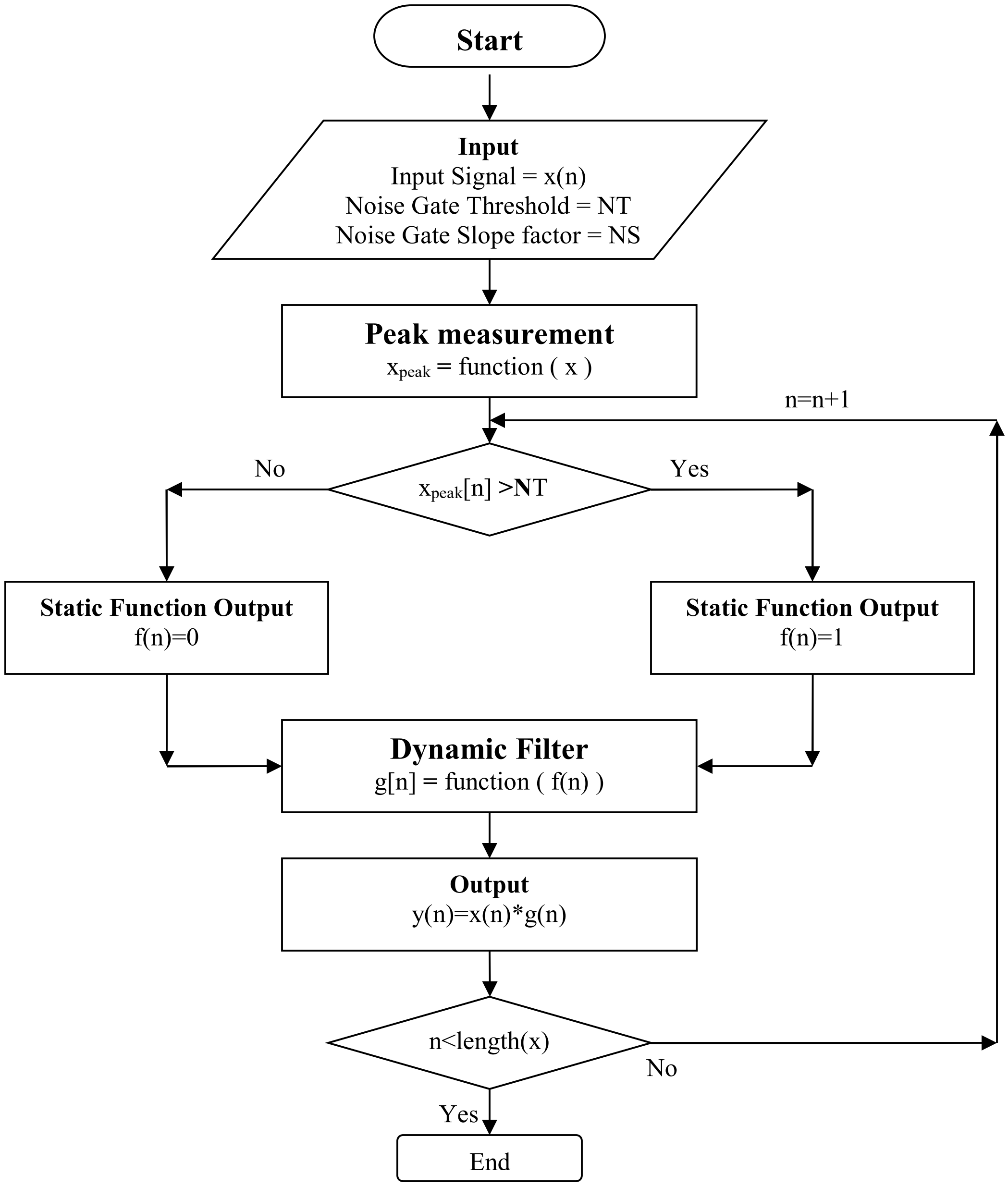
Figure 3.8 Noise Gate

Noise gating modulates the output off when the input level is below a threshold. The modulation may be a square wave, or a variation of expansion where the low level inputs are "expanded" down into silence, which gives a less abrupt transition.

The decision to activate the gate is based on a peak measurement which leads to a fade in/fade out of the gain factor g(n) with appropriate attack and release times.

The input to “time constant system” is set to zero if the input level falls below the noise gate threshold & is set to one if the input level exceeds the noise gate level.

The main use of a noise gate is to eliminate noise when the desired signal is not present. The noise gate attenuates only the soft signals.



Flow Chart 3.10 Working of a Noise

Chapter 4

Effects Based on variation of Frequency

Chapter 4
Effects Based on variation of Frequency

These are effects based on filtering the input signal or modulated of its frequency. Following are the some of these effects.

- Pitching Shifting
- Phaser
- Wah-Wah
- Vibrato

4.1. - Phasing

If two signals that are identical, but out of phase, are added together, then the result is that they will cancel each other out. If, however, they are partially out of phase, then partial cancellations and partial enhancements occur. This leads to the phasing effect.

A Phaser can be realized by cascading all-pass filters with time varying centre frequency. The output signal from these cascaded filter are added to original signal with causes phase cancellation or enhancements.

Phasing effect is very similar to flanging effect that's why it is often considered to go along delay based effect, but it does not rely on delay at all.

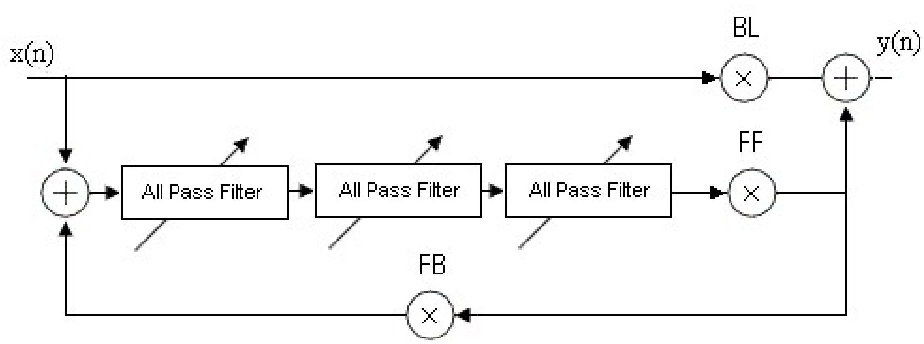


Figure 4.1: Structure for Phaser

Transfer function of All-pass filter: $A(z) = \dfrac{-c + d(1-c)z^{-1} + z^{-2}}{1 + d(1-c)z^{-1} - cz^{-2}}$ 4.1

Where $c = \dfrac{\tan(\pi f_b / f_s) - 1}{\tan(2\pi f_b / f_s) + 1}$ $\quad d = -\cos(2\pi f_c / f_s)$ 4.2

Difference Equation of All-pass filter

$y_b[n] = -cx[n] + d(1-c)x[n-1] + x[n-2] - d(1-c)y_b[n-1] + cy_b[n-2]$ 4.3

f_b = Bandwidth

f_c = Centre Frequency

f_s = Sampling Frequency

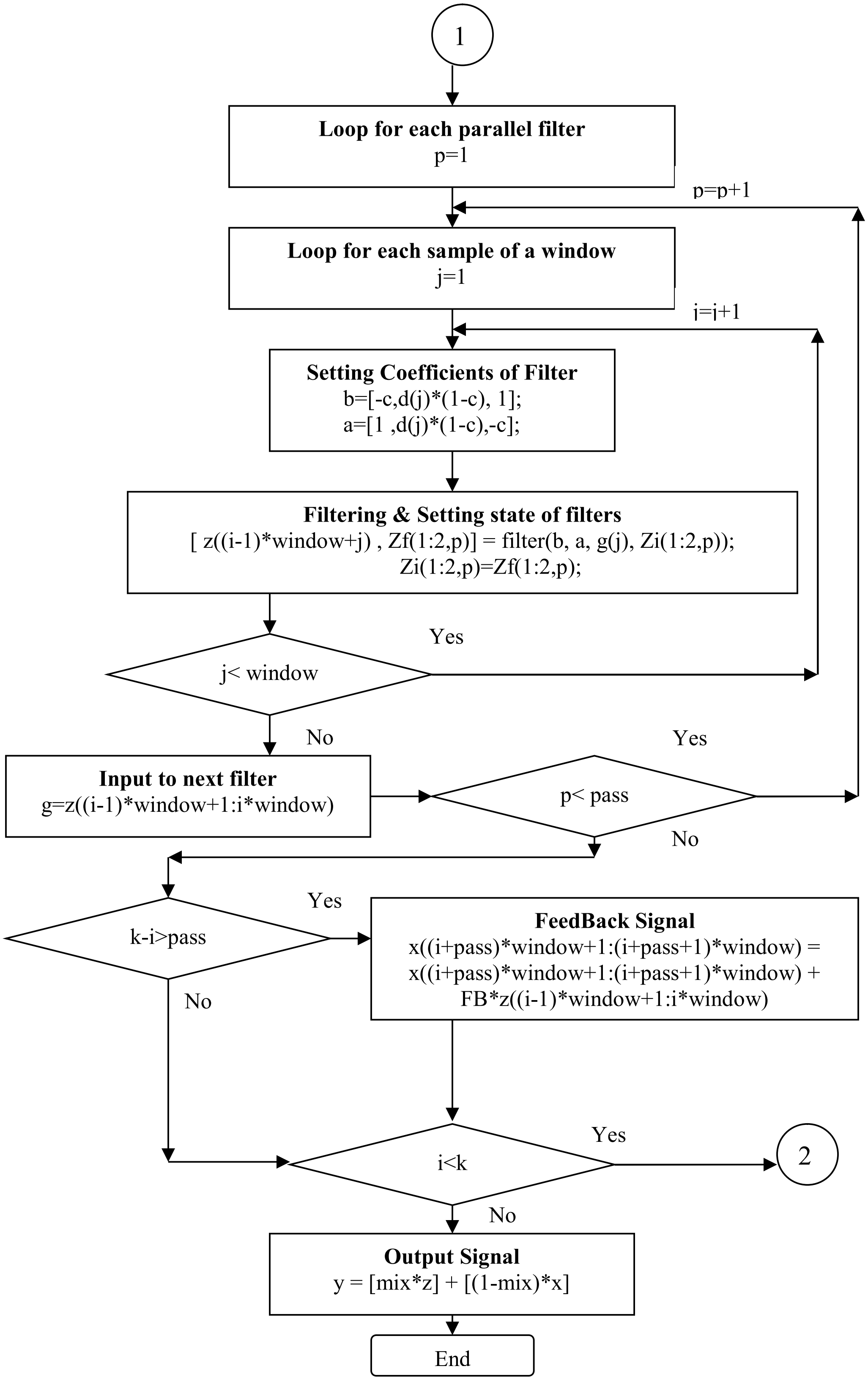


Figure 4.1 Working of "Phaser Filter" Simulator Part-2

4.2. Wah-Wah

A resonator that can have its centre frequency moved up or down in frequency by moving a pedal. The "wah" name comes from the way it mimics the moving resonance of the human vocal tract in speech as the sound "wah" is made.

If the variation of the centre frequency is controlled by the input signal a low frequency oscillator is used to change the centre frequency. Such an effect is called an auto wah filter.

If the effect is combined with a low frequency amplitude variation which produce a tremolo the effect is denoted a tremolo-wah filter.

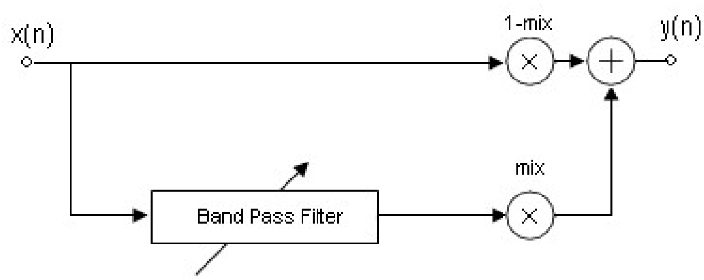


Figure 4.2 Wah-wah Filter (with single time varying Band-pass Filter)

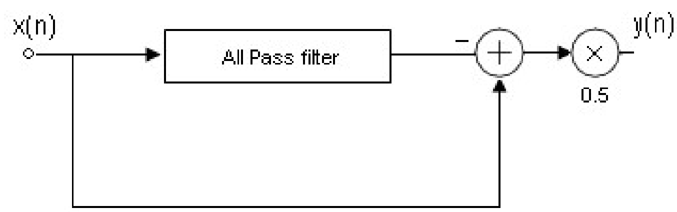


Figure 4.3 Band-pass Filter

Transfer function of All-pass filter: $A(z) = \dfrac{-c + d(1-c)z^{-1} + z^{-2}}{1 + d(1-c)z^{-1} - cz^{-2}}$ 4.4

Transfer function of BP filter: $H(z) = \dfrac{1}{2}[1 - A(z)] = \dfrac{(1+c) - (c+1)z^{-2}}{2 + 2d(1-c)z^{-1} - 2cz^{-2}}$ 4.5

Where $c = \dfrac{\tan(\pi f_b / f_s) - 1}{\tan(2\pi f_b / f_s) + 1}$ $\quad d = -\cos(2\pi f_c / f_s)$ 4.6

Difference Equation of All-pass filter

$$y_b[n] = -cx[n] + d(1-c)x[n-1] + x[n-2] - d(1-c)y_b[n-1] + cy_b[n-2] \qquad 4.7$$

Difference Equation of Band-pass filter

$$y[n] = 0.5(x[n] - y_b[n]) \qquad 4.8$$

f_b = Bandwidth

f_c = Centre Frequency

f_s = Sampling Frequency

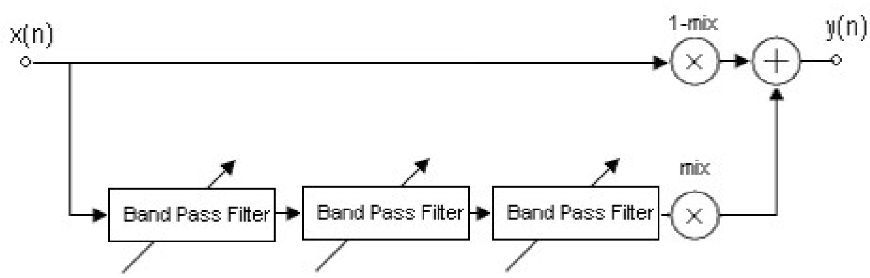


Figure 4.4 M-Fold Wah-wah [Dis99]

Replacing single band pass filter with multiple band-pass filter leads to the M-Fold wah-wah filter (shown in above figure). M band-pass filters are spread over the entire spectrum and simultaneously change their centre frequency.