IJSBAR
International Journal of Sciences: Basic and Applied Research

International Journal of
Sciences: Basic and Applied Research

AUDIO PROCESSING ANALYZER

By *Sana Rizwan*

Volume 35, 2017
ISSN (Print & Online): 2307-4531

Published by:
GSSRR
Global Society of Scientific Research and Researchers
Visit: www.gssrr.org
ISSN 2307-4531 (Print & Online)

IJSBAR research papers are currently indexed by:

AUDIO PROCESSING ANALYZER

ISSN(online & Print) 2307-4531

The IJSBAR is published and hosted by the Global Society of Scientific Research and Researchers (GSSRR).

Members of the Editorial Board

GSSRR
Global Society of Scientific Research and Researchers

IJSBAR
International Journal of Sciences: Basic and Applied Research

PROJECT THESIS

AUDIO PROCESSING ANALYZER

AUDIO PROCESSING ANALYZER

by

Sana Rizwan

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Software Engineering

COMSATS INSTITUTE OF
INFORMATION TECHNOLOGY

2012

Approved by ____
Chairperson of Supervisory Committee

____

____

____

Program Authorized
to Offer Degree ____

Date ____

COMSATS INSTITUTE OF
INFORMATION TECHNOLOGY
ABSTRACT

AUDIO PROCESSING ANALYZER

by Sana Rizwan

Chairperson of the Supervisory Co________________________
Department of Computer Engineering

The project emphasizes simulation of various DSP effects using elementary phenomenon of audio processing, and by manipulating audio using various filters in order to enhance the quality. There are many commercially available systems, which provide facilities such as channel equalizers, karaoke systems, and a few audio processors based on Digital Signal Processing. Software systems are also available which provide a fairly good and cost effective solution to audio enhancement. Yet they are limited due to resources issues and thus make a trade-off between performance and quality. The project at first studies and analyzes proceeds as study and analysis of audio processing phenomena and various effects involved in it. In the second phase algorithms have been developed for these phenomena and their simulation in MATLAB.

ACKNOWLEDGMENTS

I am indebted to my project supervisor who has contributed in the development of this report. In particular i wish to heartily thank my appreciation to my professor Madam Saba for providing me necessary insight into Digital Audio processing and guiding me on how to carry on the project successfully. I thank her for her commitment to excellence in all aspects of the production of this report. Her creativity, assistance and patience are truly appreciated. I am also grateful to thank Sir Kashif Ayub of the Computer Science Department at COMSATS, for his suggestions to improve the Graphical Interface of the project.

Table of Contents

GLOSSARY

Word. [Click and type definition here.]

REFERENCE

Chapter 1

INTRODUCTION TO THE REPORT

1.1 Thesis organization

The report is organized in 5 Chapters. Each Chapter contains its own introduction. Table of contents and abstract is provided at the start of the thesis. Bibliography and Index is provided at the end.

Chapter 1 provides the overall layout of the thesis. Chapter 2 provides the basic idea of sound and it’s propagation along with brief introduction to how humans perceive sound. Furthermore it also covers few audio effects that are widely used to enhance audio processing. Project is the implementation of different audio effects that can create different environments by changing the existing audio. Digital signal processing, which is an essential part of the project is explained in Chapter 3. Project explains a detailed overview of audio and audio effects. Effects are explained with the help of flowcharts in Chapter 4. The algorithms are also explained in the same chapter, which helps in better understanding of the audio effects. The results, after applying the effects are shown which provides a clear idea of the input and the effective output.

Outputs of the effects are shown in Chapter 4 with diagrams. Thesis is concluded by giving future recommendations in Chapter 5.

1.2 Scope and Motivation

The science of audio production and recording has benefited greatly from the development of new digital signal processing technologies. Nearly every application in music production that was once handled by analog equipment is now handled by cleaner, faster, and more reliable digital hardware. In addition to the replacement and improvement of analog hardware, fast DSP chips are now being used to process audio signals in new and amazing ways. DSP chips make applications such as advanced noise reduction, pitch estimation, and time stretching possible.

In the present time of exceedingly commercialized and industrialized music production, sound processing, post production and effects crafting have become the central part of making music. Practically no one can do without some effects and often mastering can make a whole world of difference in the world of music particularly the music industry and marketing. To satisfy these needs, an overwhelming number of effects and processing engines, algorithms, environments, plug-in and alike have become available to satisfy the various needs of the world of music. Effects algorithms are an integral part of many synthesis platforms. This means that understanding effects algorithms is vital to any electronic musician that aims at doing something useful and productive.

Chapter 2

HUMAN PERCEPTION OF SOUND

The chapter covers the basic idea of sound and its production. It begins with the definition of sound and an explanation of the human perception of sound, process and range of frequency which a human ear can recognize and goes on to explain the propagation of sound waves from one medium to the other. Later in the chapter it explains the ranges of frequencies ,which a human ear can recognize.

2.1 Sound

Sound is defined as the alteration of pressure that propagates through an elastic medium such as air, which produces an auditory sensation.

2.1.1 Production of sound

A sound wave, like any other wave, is introduced into a medium by a vibrating object. The vibrating object is the source of the disturbance which moves through the medium.[1]

The vibrating object, which creates the disturbance, could be the vocal chord of a person, the vibrating string and soundboard of a guitar or violin, the vibrating tines of a tuning fork, or the vibrating diaphragm of a radio speaker. Regardless of what vibrating object is creating the sound wave, the particle of the medium through which the sound moves vibrate in a back and forth motion at a given frequency. The frequency of a wave is measured as the number of complete back-and-forth vibrations of a particle of the medium per unit of time. If a particle of air undergoes 1000 longitudinal vibrations in 2 seconds, then the frequency of the wave would be 500 vibrations per second. A commonly used unit for frequency is the Hertz (Hz), where

$$1 \text{ Hertz} = 1 \text{ vibration/second}$$

As a sound wave moves through a medium, each particle of the medium vibrates at the same frequency. This is sensible since each particle vibrates due to the motion of its nearest neighbor. The first particle of the medium begins vibrating, at say 500 Hz, and begins to set the second particle into vibrational motion at the same frequency i.e 500 Hz. The second particle begins vibrating at 500 Hz, it sets the third particle of the medium into vibrational motion at 500 Hz. The process continues throughout the medium; each particle vibrates at the same frequency. And of course the frequency at which each particle vibrates is the same as the frequency of the original source of the sound wave. Thus, a guitar string vibrating at 500 Hz will set the air particles in the room vibrating at the same frequency of 500 Hz carrying sound *signal* to the ear of a listener at a 500 Hz sound wave.

The back-and-forth vibrational motion of the particles of the medium would not be the only observable phenomenon occurring at a given frequency. Since a sound wave is a pressure wave, a detector could be used to detect oscillations in pressure from a high pressure to a low pressure and back to a high pressure. As the compression (high pressure) and rarefaction (low pressure) disturbances move through the medium, they would reach the detector at a given frequency. For example, a compression would reach the detector 500 times per second if the frequency of the wave were 500 Hz. Similarly, a rarefaction would reach the detector 500 times per second if the frequency of the wave were 500 Hz. Thus the frequency of a sound wave not only refers to the number of back-and-forth vibrations of the particles per unit of time, but also refers to the number of compression or rarefaction disturbances, which pass a given point per unit of time.

2.2 Human perception of sound

The ears of humans (and other animals) are sensitive detectors capable of detecting the fluctuations in air pressure, which impinge upon the eardrum. The human ear is capable of detecting sound waves with a wide range of frequencies, ranging between approximately 20 Hz to 20 000 Hz. Any sound with a frequency below the audible range of hearing (i.e., less than 20 Hz) is known as an infrasound and any sound with a frequency above the audible range of hearing (i.e., more than 20 000 Hz) is known as an ultrasound. Humans are not alone in their ability to detect a wide range of frequencies. Dogs can detect frequencies as low as approximately 50 Hz and as high as 45 000 Hz. Cats can detect frequencies as low as approximately 45 Hz and as high as 85 000 Hz. Bats, who are essentially blind and must rely on sound echolation for navigation and hunting, can detect frequencies as high as 120 000 Hz. Dolphins can detect frequencies as high as 200 000 Hz. While dogs, cats, bats, and dolphins have an unusual ability to detect ultrasound, an elephant possesses the unusual ability to detect infrasound, having an audible range from approximately 5 Hz to approximately 10 000 Hz.

The sensations of these frequencies are commonly referred to as the pitch of a sound. A high pitch sound corresponds to a high frequency and a low pitch sound corresponds to a low frequency. Many people, especially those who have been musically trained, are capable of detecting a difference in frequency between two separate sounds, which is as little as 2 Hz. When two sounds with a frequency difference of greater than 7 Hz are played simultaneously, most people are capable of detecting the presence of a complex wave pattern resulting from the interference and superposition of the two sound waves. Certain sound waves when played (and heard) simultaneously will produce a particularly pleasant sensation when heard and are said to be consonant. Such sound waves form the basis of intervals in music. For example, any two sounds whose frequencies make a 2:1 ratio are said to be separated by an octave and result in a particularly pleasing sensation when heard; that is, two sound waves sound good when played together if one sound has twice the frequency of the other. Similarly two sounds with a frequency ratio of 5:4 are said to be separated by an interval of a third; such sound waves also sound good when played together. Examples of other sound wave intervals and their respective frequency ratios are listed in the table (2.1).

Interval	**Frequency Ratio**	**Examples**
Octave	2:1	512 Hz and 256 Hz
Third	5:4	320 Hz and 256 Hz
Fourth	4:3	342 Hz and 256 Hz
Fifth	3:2	384 Hz and 256 Hz

Fig 2.1

The ability of humans to perceive pitch is associated with the frequency of the sound wave, which impinges upon the ear. Because sound waves are longitudinal waves, which produce high- and low-pressure disturbances of the particles of a medium at a given frequency, the ear has an ability to detect such frequencies and associate them with the pitch of the sound. But pitch is not the only property of a sound wave detectable by the human ear [1].

2.3 Human Hearing

The human ear is an exceedingly complex organ. Human hearing is a complex subject involving the fields of physiology, psychology and acoustics. Human ear serves as an astounding transducer, converting sound energy to mechanical energy to a nerve impulse, which is transmitted to the brain. The ear's ability to do this allows us to perceive the pitch of sounds by detection of the wave's frequencies, the loudness of sound by detection of the wave's amplitude and the timbre of the sound by the detection of the various frequencies, which make up a complex sound wave. The ear consists of three basic parts - the outer ear, the middle ear, and the inner ear. Each part of the ear serves a specific purpose in the task of detecting and interpreting sound. The outer ear serves to collect and channel sound to the middle ear. The middle ear serves to transform the energy of a sound wave into the internal vibrations of the bone structure of the middle ear and ultimately transform these vibrations into a compression wave in the inner ear. The inner ear serves to transform the energy of a compression wave within the inner ear fluid into nerve impulses, which can be transmitted to the brain.

The outer ear consists of an earflap and an approximately 2-cm long ear canal. The earflap provides protection for the middle ear in order to prevent damage to the eardrum. The outer ear also channels sound waves which reach the ear through the ear canal to the eardrum of the middle ear. Because of the length of the ear canal, it is capable of amplifying sounds with frequencies of approximately 1000 Hz to 5000Hz. As sound travels through the outer ear, the sound is still in the form of a pressure wave, with an alternating pattern of high and low pressure regions. It is not until the sound reaches the eardrum at the interface of the outer and the middle ear that the energy of the mechanical wave becomes converted into vibrations of the inner bone structure of the ear.

The middle ear is an air-filled cavity, which consists of an eardrum and three tiny, interconnected bones - the hammer, anvil, and stirrup. The eardrum is a very durable and tightly stretched membrane, which vibrates as the incoming pressure waves reach it. Compression forces the eardrum inward and a rarefaction forces the eardrum outward, thus vibrating the eardrum at the same frequency of the sound wave. Being connected to the hammer, the movements of the eardrum will set the hammer, anvil, and stirrup into motion at the same frequency of the sound wave. The stirrup is connected to the inner ear; and thus the vibrations of the stirrup are transmitted to the fluid of the middle ear and create a compression wave within the fluid. The three tiny bones of the middle ear act as levers to amplify the vibrations of the sound wave. Due to a mechanical advantage, the displacements of the stirrup are greater than that of the hammer. Since the pressure wave striking the large area of the eardrum is concentrated into the smaller area of the stirrup, the force of the vibrating stirrup is nearly 15 times larger than that of the eardrum. This feature enhances the ability to hear the faintest of sounds. The middle ear is an air-filled cavity, which is connected by the Eustachian tube to the mouth. This connection allows for the equalization of pressure within the air-filled cavities of the ear. When this tube becomes clogged during a cold, the ear cavity is unable to equalize its pressure; this will often lead to earaches and other pains.

The inner ear consists of a cochlea, the semicircular canals, and the auditory nerve. The cochlea and the semicircular canals are filled with a water-like fluid. The fluid and nerve cells of the semicircular canals provide no roll in the task of hearing; they merely serve as *accelerometers* for detecting accelerated movements and assisting in the task of maintaining balance. The cochlea is a snail-shaped organ, which would stretch to approximately 3 cm. In addition to being filled with fluid, the inner surface of the cochlea is lined with over 20 000 hair-like nerve cells which perform one of the most critical roles in our ability to hear. These nerve cells have a differ in length by minuscule amounts; they also have different degrees of resiliency to the fluid which passes over them. As a compressional wave moves from the interface between the hammer of the middle ear and the *oval window* of the inner ear through the cochlea, the small hair-like nerve cells will be set in motion. Each hair cell has a natural sensitivity to a particular frequency of vibration. When the frequency of the compression wave matches the natural frequency of the nerve cell, that nerve cell will resonate with larger amplitude of vibration. This increased vibrational amplitude induces the cell to release an electrical impulse, which passes along the auditory nerve towards the brain. In a process, which is not clearly understood, the brain is capable of interpreting the qualities of the sound upon reception of these electric nerve impulses.

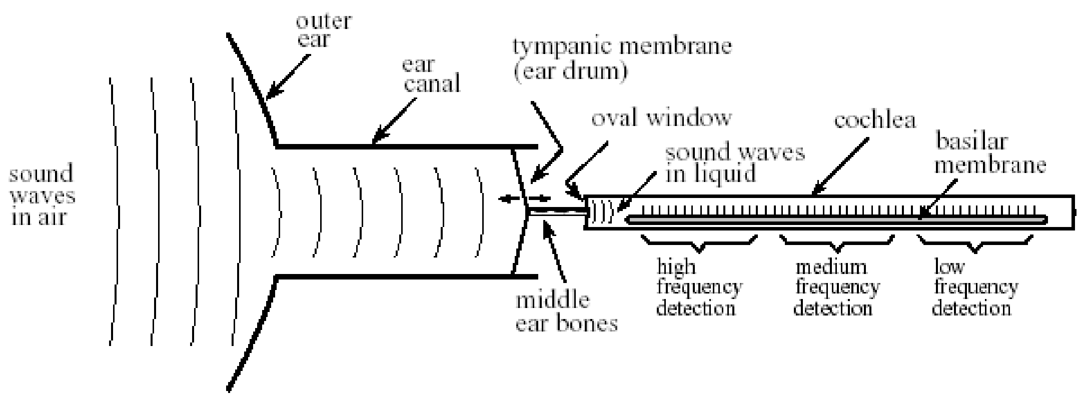


Fig 2.3. The outer ear collects sound waves from the environment and channels them to the tympanic membrane (ear drum), a thin sheet of tissue that vibrates in synchronization with the air waveform. The middle ear bones (hammer, anvil and stirrup) transmit these vibrations to the oval window, a flexible membrane in the fluid filled cochlea. Contained within the cochlea is the basilar membrane, the supporting structure for about 12,000 nerve cells that form the cochlear nerve. Due to the varying stiffness of the basilar membrane, each nerve cell only responses to a narrow range of audio frequencies, making the ear a frequency spectrum analyzer.

A nerve cell on the basilar membrane can encode audio information by producing an action potential in response to each cycle of the vibration. For example, a 200 hertz sound wave can be represented by a neuron producing 200 action potentials per second. However, this only works at frequencies below about 500 hertz, the maximum rate that neurons can produce action potentials. The human ear overcomes this problem by allowing several nerve cells to take turns performing this single task.

It is common to express sound intensity on a logarithmic scale, called decibel SPL (Sound Power Level). On this scale, 0 dB SPL is a sound wave power of 10-16 watts/cm2, about the weakest sound detectable by the human ear. Normal speech is at about 60 dB SPL, while painful damage to the ear occurs at about 140 dB SPL. The difference between the loudest and faintest sounds that humans can hear is about 120 dB, a range of one million in amplitude. Listener can detect a change in loudness when the signal is altered by about 1 dB i.e a 12% change in amplitude. There are only about 120 levels of loudness that can be perceived from the faintest whisper to the loudest thunder. When listening to very weak sounds, the eardrum vibrates less than the diameter of a single molecule. The perception of loudness relates roughly to the sound power to an exponent of 1/3. The range of human hearing is generally considered to be 20 Hz to 20 kHz, but it is far more sensitive to sounds between 1 kHz and 4 kHz. For example, listeners can detect sounds as low as 0 dB SPL at 3 kHz, but require 40 dB SPL at 100 hertz i.e. an amplitude increase of 100. Listeners can tell that two tones are different if their frequencies differ by more than about 0.3% at 3 kHz. This increases to 3% at 100 hertz. For comparison, adjacent keys on a piano differ by about 6% in frequency.

Sound intensity is expressed as power per unit area (Such as watt/cm^2), or more commonly on a logarithmic scale called decibels SPL. Human hearing is the most sensitive between I KHz and 4 KHz.

The primary advantage of having *two* ears is the ability to identify the direction of the sound. Human listeners can detect the difference between two sound sources that are placed as little as three degrees apart, about the width of a person at 10 meters. This directional information is obtained in two separate ways. First, frequencies above about 1 kHz are strongly shadowed by the head. In other words, the ear nearest the sound receives a stronger signal than the ear on the opposite side of the head. The second clue to directionality is that the ear on the far side of the head hears the sound slightly later than the near ear, due to its greater distance from the source. Based on a typical head size (about 22 cm) and the speed of sound (about 340 meters per second), an angular discrimination of three degrees requires a timing precision of about 30 microseconds. Both these sources of directional information are greatly aided by the ability to turn the head and observe the change in the signals. An interesting sensation occurs when a listener is presented with exactly the same sounds to both ears, such as listening to monaural sound through headphones. The brain concludes that the sound is coming from the center of the listener's head. While human hearing can determine the direction a sound is from, it does poorly in identifying the distance to the sound source. This is because there are few clues available in a sound wave that can provide this information. Human hearing weakly perceives that high frequency sounds are nearby, while low frequency sounds are distant. This is because sound waves dissipate their higher frequencies as they propagate long distances. Echo content is another weak clue to distance, providing a perception of the room size. For example, sounds in a large auditorium will contain echoes at about 100 millisecond intervals, while 10 milliseconds is typical for a small office.

2.4 Sound Quality and Data Rate

Audio processing covers many diverse fields, all involved in presenting sound to human listeners. When designing a digital audio system there are two questions that are important to be noticed.

- How good does it need to sound?
- What data rate can be tolerated?

The categories defined below can satisfy these questions

1. High fidelity music

Where sound quality is of the greatest importance and almost any data rate will be acceptable. High fidelity music systems sample fast enough at 44.1 kHz, and with enough precision of 16 bits, that they can capture virtually all of the sounds that humans are capable of hearing. This magnificent sound quality comes at the price of a high data rate, 44.1 kHz ×

16 bits = 706k bits/sec. Whereas music requires a bandwidth of 20 kHz, natural sounding speech only requires about 3.2 kHz. Even though the frequency range has been reduced to only 16% (3.2 kHz out of 20 kHz), the signal still contains 80% of the original sound information.

2. Telephone communication

It requires natural sounding speech and a low data rate to reduce the system cost. Telecommunication systems typically operate with a sampling rate of about 8 kHz, allowing natural sounding speech, but greatly reduced music quality. FM radio stations broadcast with a bandwidth of almost 20 kHz, while AM radio stations are limited to about 3.2 kHz. Voices sound normal on the AM stations, but the music is weak and unsatisfying. Voice-only systems also reduce the precision from 16 bits to 12 bits per sample, with little noticeable change in the sound quality. This can be reduced to only 8 bits per sample if the quantization step size is made unequal. This procedure is called 'companding'

3. Compressed speech

Where reducing the data rate is very important and some unnaturalness in the sound quality can be tolerated. An 8 kHz sampling rate, with an ADC precision of 8 bits per sample, results in a data rate of 64k bits/sec. Speech requires less than 10% of the data rate of high fidelity music. The data rate of 64k bits/sec represents the straightforward application of sampling and quantization theory to audio signals. Techniques for lowering the data rate further are based on *compressing* the data stream by removing the inherent redundancies in speech signals. . Table (2.1) shows the tradeoff between sound quality and data rate for these three categories

Sound Quality Required	Band Width	Sampling Rate	Number of bits	Data rate	Comments
High fidelity music	5Hz-20KHz	44.1KHz	16 bit	706K	Satisfies even the most picky audiophile, better than human hearing
Telephone quality speech (with companding)	200Hz-3.2Khz 200Hz-3.2KHz	8KHz 8KHz	12 bit 8 bit	96K 64K	Good speech quality, but very poor for music Nonlinear ADC reduces the data rate by 50%. A very common technique
Speech encoded by linear predictive coding	200HZ-3.2KHz	8KHz	12 bit	4K	DSP speech compression technique. Very low date rate voice quality

Fig 2.1 Audio data rate Vs Sound quality. The sound quality of a digitized audio signal depends on its data rate, the product of its sampling rate and the number of bits per sample.

Chapter 3

DIGITAL SIGNAL PROCESSING

The chapter explains the features and characteristics of DSP. Digital signal processing has been widely used. Music industry has benefited greatly from its evolution. Digital signal processing has contributed widely in Digital audio processing. Furthermore different audio effects in generalized terms are explained. Use of filters made it possible to enhance the quality of an audio and thus different types of filters are also briefed.

3.1 Digital Signal Processing (DSP):

Digital signal processing is the processing of signals by digital means.

Digital:

The technical meanings of Digital are ‘operating by the use of discrete signals to represent data in the form of numbers’.

Signal

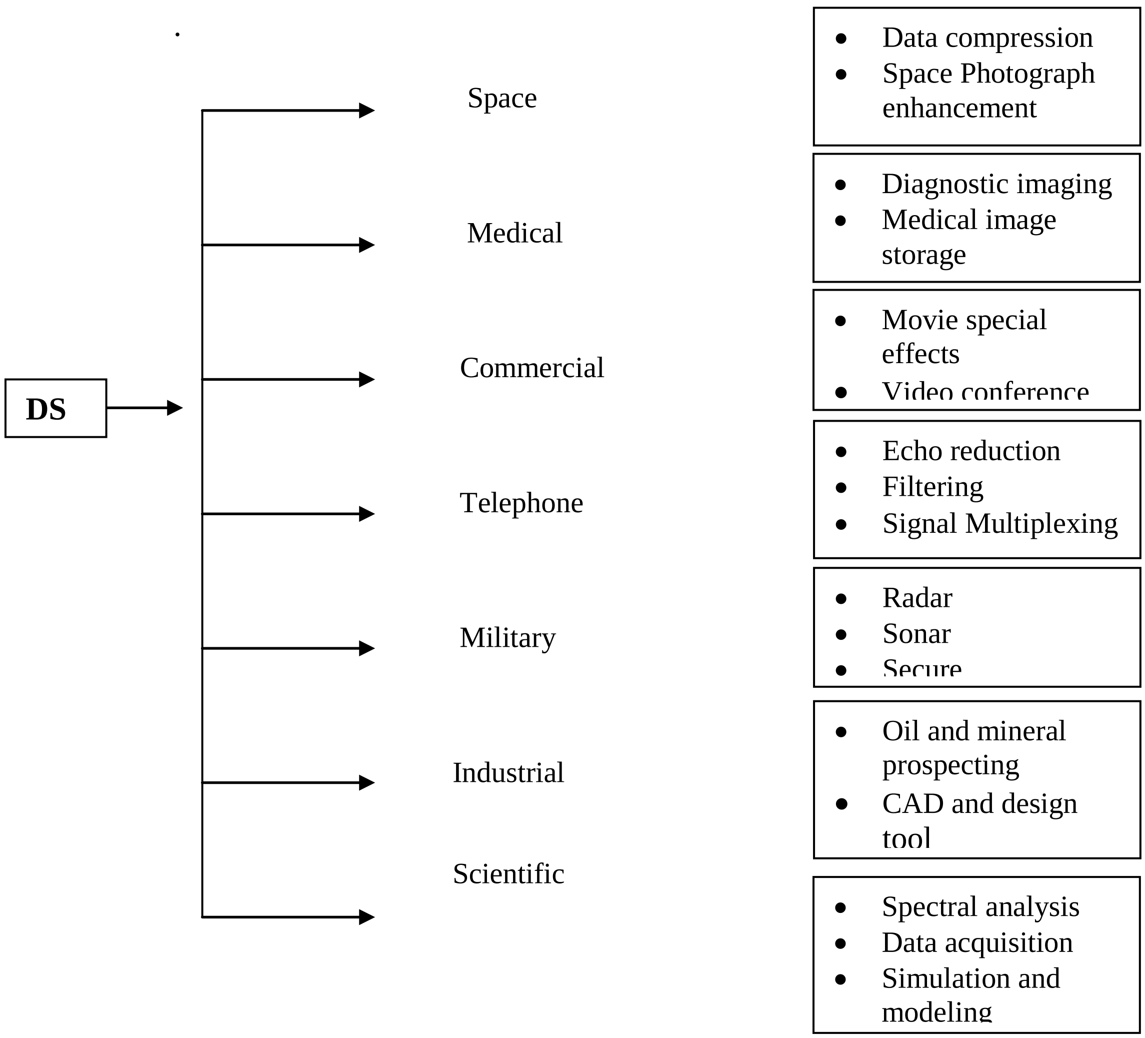
Signal in digital terminology is defined as ‘a variable parameter by which information is conveyed through an electronic circuit’.

Processing

To perform operations on data according to programmed instructions.

Digital Signal processing can also be define as the ‘Changing or analyzing information, which is, measured as discrete sequences of numbers’. Digital Signal Processing is one of the most powerful technologies that will shape science and Engineering in the twenty-first century.[4]

Communication, Medical imaging, High fidelity music reproduction. are few of the areas that develop a deep DSP technology, with its own algorithms, mathematics, and specialized techniques. This combination of breadth and depth makes it impossible for any one individual to master all of the DSP technology that has been developed DSP education which involves two tasks: learning general concepts that apply to the field as a whole, learning specialized techniques for particular area of interest



3.1.1 Development of DSP

The development of Digital Signal Processing (DSP) starts from the 1960's with the use of mainframe digital computers for number – crunching applications such as Fast Fourier Transform (FFT), which allows the frequency spectrum of a signal to be computed rapidly. These techniques were not widely used at that time, because suitable computing equipment was available only in universities and other scientific research institutions.

3.1.2 Statistics and probability

Statistics and probability are used in Digital Signal Processing to characterize signals and the processes that generate them. For example, a primary use of DSP is to reduced interference, noise, and other undesirable components in acquired data. These may be an inherent part of the signal being measured, arise from imperfections in the data acquisition system, or be introduced as an unavoidable byproduct of some DSP operation. Statistics and probability allow these disruptive features to measured and classified.

3.1.3 Conversion of Analogue Signals

Most of the signals directly encountered in science and engineering are continuous: conversion (ADC) and Digital-to-analogue (DAC) are the processes that allow digital computer to interact with these continuous signals. In this conversion information is lost due to

i. Inaccuracies in the measurement
ii. Uncertainty in timing
iii. Limits on the duration of the measurement and these effects are called quantization errors

The continuous analogue signal has to be held before it can be sampled. Otherwise the signal would be changing during the measurement .The sampling results in a discrete set of digital numbers that represent measurements of the signal –usually taken at equal intervals of time.

3.1.4 DSP Software

DSP application is usually programmed in the same languages as other science and engineering tasks, such as: C, BASIC and assembly. The power and versatile of C makes it the language of choice for computer scientists and other professional programmers. On the other hand, the simplicity of BASIC makes it ideal for scientists and engineers.

3.1.5 Signal Processing:

Signals are processed in a variety of ways. For example, the output signal from a transducer may well be contaminated with unwanted electrical 'noise'. The electrodes attached to a patient's chest when an ECG is taken measure tiny electrical voltage changes due to activity of the heart and muscles. Processing a signal using a filter can remove or at least reduce the unwanted part of the signal. Filtering of signals to improve signal quality or to extract important information is done by DSP techniques.

3.1.6 Convolution

Convolution is a mathematical way of combining two signals to form a third signal. It is the single most important technique in Digital Signal Processing. Convolution is important because it relates three signals.

- Input signal
- Output signal
- And the impulse response

3.1.7 Fundamentals

DSP system has three fundamentals sources of limitation:

i. Loss of information because we only take samples of the signal at intervals.
ii. Loss of information because we only sample the signal for a certain length of time.
iii. Errors due to limited precision (i.e. word length) in data storage and arithmetic.

The effects of these limitations are as follows:

1. Aliasing

Aliasing is produced due to the result of sampling. In aliasing the high and low frequencies are not distinguished. This cannot be overcome, because it is the limitation of any sampled data system, not just digital ones.

2. Frequency resolution

Frequency resolution is the result of limited duration of sampling; adjacent frequencies are not distinguished between each other.

3. Quantization error

It is produced due to the result of limited precision (word length) when converting between analogue and digital form, when storing data, or when performing arithmetic.

3.1.7 Features of DSP

i. Signal is discrete –which means the information in between discrete samples is lost.

ii. Signal come from the real world – this intimate connection with the real world leads to many unique needs such as the need to react in real time and a need to measure signals and convert them to digital numbers.

3.2 Applications of DSP:

DSP technology is now a day's common place in such devices as mobile phones, multimedia computers, video recorders, CD players, hard disc drive controllers and modems, and will soon replace analog circuitry in TV sets and telephones. An important application of DSP is in signal compression and decompression. In CD systems, for example, the music recorded on the CD is in a compressed form(to increase storage capacity) and must be decompressed for recorded signal to be reproduced .Signal compression is used in digital cellular phones to allow a greater number of calls to be handled simultaneously within each local “cell”. DSP signal compression technology allows people not only to talk to one another by telephone but also to see one another on the screen of their PCs, using small video cameras mounted on the computer monitors, with only a conventional telephone line linking them together.

Although the mathematical theory underlying DSP techniques such as Fast Fourier Transform, digital filter design and signal compression can be fairly complex, the numerical operations required to implement these techniques are in fact very simple, consisting mainly of operations that could be done on a cheap four function calculator. The architecture of a DSP chip is designed to carry out such operations incredibly fast, processing up to tens of millions of samples per second, to provide *real time* performance: that is ability to process a signal “live” as it is sampled and then output the processed signal, for example to a loudspeaker or video display. All of the practical examples of DSP applications mentioned earlier, such as hard disc drives and mobile phones, demand real –time operation.

3.3 Digital Signal Processors (DSPs)

The introduction of microprocessors in the late 1970's and early 1980's made it possible for DSP techniques to be used in much wider range of applications. General – purpose microprocessors such as the Intel x86 family are not ideally suited to the numerically- intensive requirements of DSP, and during the 1980's the increasing importance of DSP led several major electronics manufacturers to develop Digital Signal Processor chips – specialized microprocessors with architecture designed specifically for the types of operations required in digital signal processing.

DSP is a programmable device, with its own native instruction code. DSP chips are capable of carrying out millions of floating point operations per second, and like better-known general-purpose cousins, faster and more powerful versions are continually being introduced.

3.4 Advantages of DSP

1. Versatility

i. Digital systems can be reprogrammed for other applications (at least where programmable DSP chips are used)

ii. digital systems can be ported to different hardware (for example a different DSP chip or board level product)

2. Repeatability

i. Digital Systems can be easily duplicated.

ii. Digital Systems do not depend on strict component tolerances.

iii. Digital System responses do not drift with temperature.

3. Simplicity

i. some things can be done more easily digitally than with analogue systems

3.5 DSP in Audio Processing:

The two principal human senses are vision and hearing. Correspondingly, much of DSP is related to image and audio processing. DSP has made revolutionary changes in both these areas.

3.5.1 Audio Processing

Audio processing covers a wide range of means and techniques involved in presenting sound to human listeners. With the new advancements and research in the field of audio and signals, there rose a need to integrate the two acquire breakthrough in the field of audio. Audio processing has been a part of life now over half a century, when music and other audio applications were being introduced. Audio Processing includes various ideas, like production of good quality music, noise cancellation, and some other effects, which absolutely change the feel of the audio signal and its quality. There are some real-life phenomena, which are sometimes; very useful in producing and using different algorithms on audio signals using the Digital Signal Processing techniques. Through these techniques one can simulate special effects or are necessary for environment creation. These real-life phenomena include time-based effects like Echo, Reverberation and Flange. These are simple to simulate on an audio signal by introducing and manipulating delays on the signals.

There are some other effects, which employ frequency for their operation, and change the frequency contents of the signals. These are called frequency-oriented or Quality Augmentation effects. These include various filters, and their integration with some amplitude-based effects. These are used for many purposes like Noise cancellation, voice identification and many other applications.

3.5.2 Music

The path leading from the musician's microphone to the audiophile's speaker is remarkably long. Digital data representation is important to prevent the degradation commonly associated with analog storage and manipulation. This is very familiar to anyone who has compared the musical quality of cassette tapes with compact disk. In a typical scenario, a musical piece is recorded in a sound studio on multiple channels or tracks. In some cases, this even involves recording individual's instruments and singers separately. This is done to give the sound engineer greater flexibility in creating the final product. The complex process of combining the individual tracks into a final product is called *mix down*. DSP can provide several important functions during mix down, including: filtering, signal addition and subtraction, signal editing, etc.

One of the most interesting DSP applications in music preparation is *artificial reverberation.* If the individual's channels are simply added together, the resulting piece sounds frails and diluted, much as if the musicians were playing outdoors. This is because listeners are greatly influenced by the echo or reverberation content of the music, which is usually minimized in the sound studio. DSP allows artificial echoes and reverberation to be added during mix down to simulate various ideal listening environments. Echoes with delays of a few hundred milliseconds give the impression of cathedral like locations. Adding echoes with delays of 10-20 milliseconds provide the perception of more modest size listening rooms.

Audio Effects

There are three categories of effects, which can be applied and synthesized with audio

1. Amplitude-based effects
2. Real-life phenomena/time-oriented/Delay-based effects
3. Quality augmentation/frequency-based effects

3.6.1 Amplitude-based effects

These effects include the following:

i. Volume Control

ii. Balance

iii. Phasing

iv. Expansion

v. Noise gating

1. Volume control

Volume control is the controlling of the amplitude of the signal by varying the attenuation of the input signal. An Active volume control will have the ability to increase the volume (i.e. amplify the input signal) as well as attenuating the signal.

Volume controls are useful for placing between effects so that the relative volume of the different effects can be kept at the constant level if not all effects have volume control built-in, allowing the user to adjust the volume of the output with effect on relative to the volume of the unaffected signal.[5]

2. Balance

Balance is used in stereo recordings. Stereo recording have two channels can be adjusted – this adjustment effectively adjusts the position of the perceived sound within the stereo field. The two extreme being: all sounds completely on the left, or all sound completely on the right. This is commonly referred to as *balance* on commercial systems. Balance is added to the stereo effect, but it does not help with stereo separation.[5]

3. Phasing

Phasing is another effect that comes under the amplitude category. Phasing is something that carries a key significance in some of the other advanced effects and noise reduction techniques. Phasing actually is inverting the phase of the audio signal, this is helpful while canceling noise from the audio signal, The technique is to find out the frequency of noise in the signal, filter out the noise signal from any one channel, apply phasing to it, and add it to the other channel. Similar is the way for an effect called Karaoke, which removes the vocals from a song.

4. Expansion

The expander is a type of dynamic processor. It increases the dynamic range of a signal such that low level signals are attenuated while the louder portions are neither attenuated nor amplified.[5]

5. Noise gating

A noise gate, quite simply, gates (or blocks) signals whose amplitude lies below a certain threshold, and lets other signals through. This is useful for eliminating background noises, such as hiss or hum, during periods of silence in a recording or performance. At other times, the recording or performance usually drowns out the background noise.

Noise gates usually have controls for hold time, attack time, and release time**.** The hold time is the time for which a signal should remain below the threshold, before it is gated. The attack time is the time during which a signal (that is greater than the threshold) is faded in from the gated state. The release time is the time during which a signal (that is below the threshold) is faded into the gated state. These controls help to eliminate the problems of distortion caused by gating signals that are part of the foreground audio signal, and the problem of sustained notes being suddenly killed by the noise gate.[5]

3.6.2 Delay-based effects

These effects include some of the real-life phenomena, which are sometimes required to be applied to audio, in order to simulate the same, or introduce some kind of an enhancement by applying any one some of them to the audio signal, virtually creating an environment in the audio. These effects hold their importance in the music industry. The effects under this category are as follows

1. Flange
2. Echo
3. Reverberation
4. Chorus

1. Flange

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 flange is varied between 0ms and 5ms 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. Flanging gives a "whooshing" sound, like the sound is pulsating. It is essentially an exaggerated chorus.[5]

2. Echo

Echo is produced by adding a time-delayed signal to the output. This produces a single echo. Multiple echoes are achieved by feeding the output of the echo unit back into it's input through and attenuator. The attenuator determines the **decay** of the echoes, which is how quickly each echo dies out. Echo greatly improves the sound of a distorted lead guitar solo, because it improves the sustain and gives an overall smoother sound. Very short echoes (5 to 15ms) 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.[5]

3. Reverberation

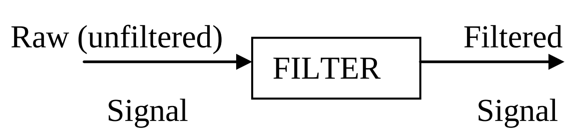
Reverberation (reverb for short) is probably one of the most heavily used effects in music. Reverb is used to simulate the acoustical effect of rooms and enclosed buildings. In a room, for instance, sound is reflected off the walls, the ceiling and the floor. The sound heard at any given time is the 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 it's original magnitude.[5]

4. Chorus

The chorus effect is so named because it makes the recording of a vocal track sound like two or more people signing in chorus sang it. This is achieved by adding a single delayed signal (echo) to the original input. However the delay of this echo is varied continuously between a **minimum delay** and a **maximum delay** at a certain rate. Typically the delay is varied between 40ms and 60 ms at a rate of about 0.2Hz.[5]

3.7 Filter

Filter is used to remove unwanted parts of signal, such as random noise, or to extract useful parts of the signal, such as the components are lying within a certain range of frequency range.



There are two main kinds of filter, analog and digital. They are quite different in their physical makeup and in how they work. Analog filter uses analog electronic circuits made up from components such as resistors, capacitors and op amps to produce the required filtering effect. Such filter circuits are widely used in such applications as noise reduction, Video signal enhancement, and graphic equalizes in hi-fi systems, and many other areas. There are well establishing techniques for designing an analog filter circuit for a given requirement. At all stages, the signal being filtered in an electric voltage or current, which is the direct analogue of the physical quantity (e. g a sound or video signal transducer output) involved.

A digital filter uses a digital processor to perform numerical calculations such as a PC, or a specialized DSP (Digital Signal Processors) chip. In a digital filter a signal is represented by a sequence of numbers, rather than a voltage or current.

3.7.1 Categories of filter

1. Non Recursive or FIR (Finite Impulse Response)
2. Recursive or IIR (Infinite Impulse Response)

1. Non-Recursive or FIR (Finite Impulse Response)

An FIR filter is one whose impulse response is of finite duration, because the current output is calculated solely from the current and previous input values.

2. Order of Non recursive (FIR) filter

The order of non-recursive: (FIR) filter is merely the number of taps used in the filter structure.

3.7.2 Design of FIR filter

Three types of methods are used to design FIR filters.

1. Parks-McClellen

It is an iteration algorithm that accepts filter specifications in terms of pass band and stop band frequencies, pass band ripple, and stop band attenuation. The PM method cannot only design FIR “filters” but also FIR “differentiators” and FIR “Hilbert transformers”.

2. Windowing

In this method taking the Inverse Discrete Fourier Transform (IDTF) of the desired frequency response derives an initial impulse response. Then, applying a data window to it refines the impulse response.

3. Direct Calculation

The impulse responses of certain types of FIR filters can be calculated directly by formulas.

3.7.3 Implementation of FIR filters:

Structurally, FIR filters consist of just two things: a sample delay line and a set of coefficients. To implement the filter:

1. Put the input sample into the delay line

2. Multiply each sample in the delay line by the corresponding coefficient and accumulate the result.
3. Shift the delay line by one sample to make room for the next input sample.

3.7.4 Advantages of FIR filter

- They can easily be designed to be “linear phase”.
- They are simple to implement. On most microprocessors, looping a single instruction can do the FIR calculation.
- They are suited to multi-rate applications. Multi-rate meaning
  by Decimation (reducing the sampling rate)
  by interpolation (increasing the sampling rate) or
  by both.
- They can be implemented using fractional arithmetic.

3.8 Recursive or IIR (Infinite Impulse Response)

An IIR filter is one who’s impulse response in infinite or (theoretically) continues for ever, because the recursive (previous output) terms feed back energy into the filter input and keeps it going.

3.8.1 Order of Recursive (IIR) filter

The order of recursive filter is the largest number of previous input or output values required to compute the current output.

3.8.2 Types of filter

1. High pass filter
2. Low pass filter
3. Band pass filter

1. High pass filter

This is a kind of a filter in which a transmitting band starts at a lower cutoff frequency and extending to (theoretically) infinite frequency. There are many varied ways of implementing a high pass filter. The simplest way is to take a pixel and subtract it from its neighbors. If the pixel is in an area of little change, such as the middle of the box, then the difference between the pixel and its neighbors will be zero. However if the pixel is on the edge of the box then the difference will be large.

2. Low pass filter

The filter passes all frequencies below a specified frequency with a little or no loss, but strongly attenuates higher frequencies. A high quality Lowpass filter should be look more like the “box car” amplitude response. It is impossible to achieve the ideal response exactly using a finite order filter, for this purpose to achieve the improved amplitude response, add more poles and zeros in the implementation.

The classical methods to design Lowpass filter are derived from analog Butterworth,Chebyshev, and Elliptic Function filters for designing the Lowpass filter Butterworth filters are optimal in sense of having a maximally flat amplitude response.

3. Band pass filter

A Bandpass filter is usually a lowpass and high pass filter in series, allowing only a certain range of frequencies through. Because the cut-off frequencies are close to one another, the effect will be similar to that of a peaking filter.

3.8.3 Advantages of digital filters

1. A digital filter is programmable; a program stored in the processor's memory determines i.e. its operation. This means the digital filter can easily be changed without affecting the circuitry (hardware).
2. Digital filters can easily be designed, tested and implemented on a general-purpose computer or workstation.
3. The characteristics of analog filters circuits (particularly those containing active components) are subject to drift and are dependent on temperature. Digital filters do not suffer from these problems, and so are extremely stable with respect both time and temperature.

4. Digital signals can handle low frequency signals accurately. As the speed of DSP technology continues to increase, digital filters are being applied to high frequency signals in the RF (radio frequency) domain.
5. Digital filters are very much more versatile in their ability to process signals in a variety of ways; this includes the ability of some types of digital filter to adapt to changes in the characteristics of the signal.

Chapter 4

AUDIO EFFECTS

Development of algorithm is a very critical part while developing a software, therefore keeping this point in mind, development has been done so that there is a minimum chance of an error. The chapter covers the flow charts and the algorithms with a brief description of the effects explained which we have implemented.

CATEGORIES OF EFFECTS:

1 Amplitude-based effects

2 Real-life phenomena/time-oriented/Delay-based effects

3 Quality augmentation/frequency-based effects

4.1: Amplitude-based effects

These effects include the following:

- Volume Control
- Panning/Balance
- Phasing
- Expansion
- Noise gating

1: Volume Control

Volume is the strength of the audio signal. It is clearly the amplitude of the signal, while talking about digital signals; it is the amplitude of the samples of the audio signal. When the sound is manipulated using some system for this purpose, the input signal is considered to have its original volume. In order to increase or decrease the volume, the amplitude of the whole signal must be multiplied by some factor, which determines the increase or decrease in the volume of the signal.

When the volume of the signal is to be increased, the factor to be multiplied with the signal should be greater than 1, and less than 1 if the volume is to be decreased. A better idea of obtaining this factor can be initializing the value of volume to 1. Then add the percentage increase or decrease to it in order to get the value to be multiplied to the whole signal. Consider the following example:

Suppose the volume of an incoming signal is to be increased any 50 percent; the above procedure can be followed as:

The initial value: Volume = 1

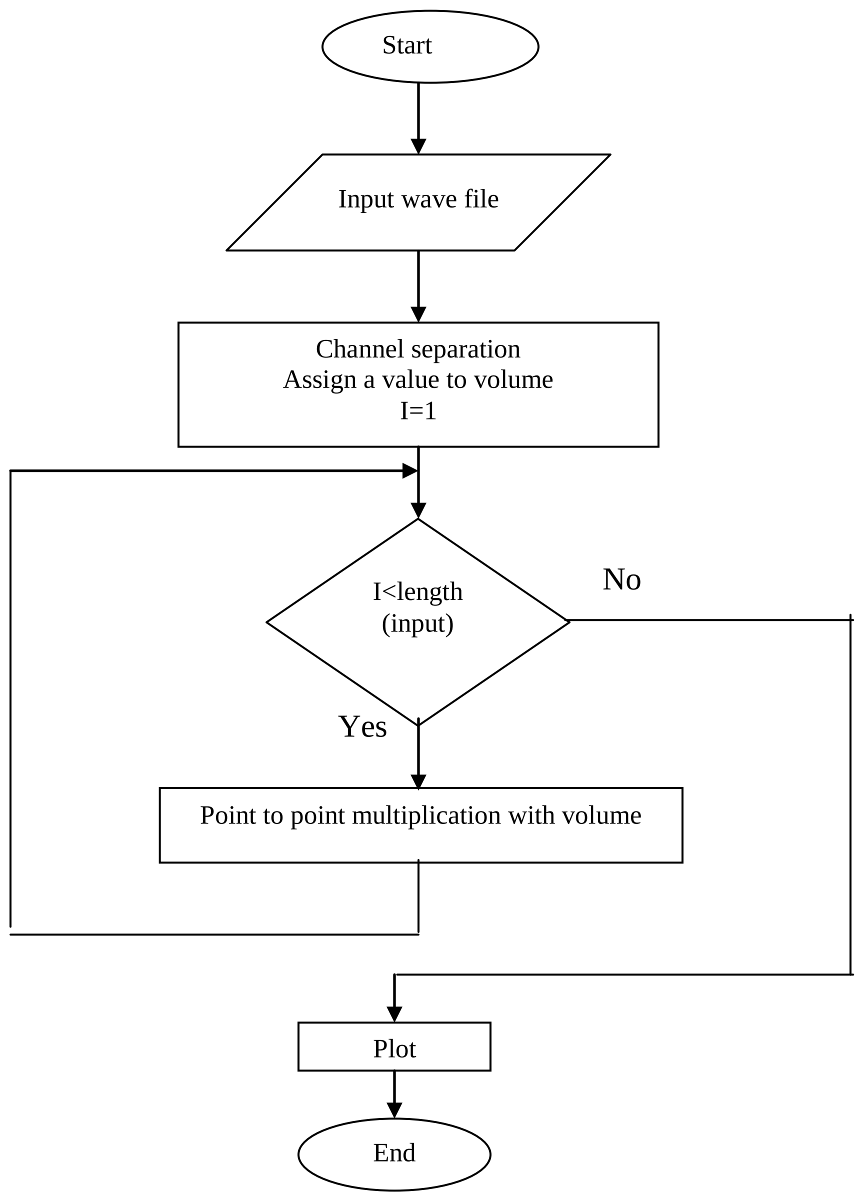
The increased value: volume = Volume + Volume*.5

(4.1)

This gives the resultant value of volume as 1.5, which amplifies the signal by 50 percent, when multiplied to the whole signal sample by sample. Similarly in case the volume is to be decreased, the percentage decrease, which is multiplied to the volume in equation (4.1), has the negative sign, thus decreasing the resultant value of the volume to be multiplied, to less than 1, and as a result, de-amplifying the whole signal by the percentage decrease.

One important thing, which should be kept in mind, is that there is strictly no sample addition in the process of volume changing, as it is only concerned to multiplication. Addition can cause signal mixing, thus contaminating the signal with unwanted values.

Volume Control



2: Panning/Balance

Panning or balance is also an amplitude-based effect, applied only to bi or multi-channel audio signals. Panning is the shifting of the intensity of the signal from one channel to another. This effect is quite useful for generating 3D surround sound or simulating the motion of the sound source. Say panning is to be applied to a two channel audio stream. The principle is the same as volume control. The only difference being that in volume control all channels are either amplified or de-amplified, while in panning, one channel is amplified and with the same percentage there is a gain in the amplitude of the other channel. Say pan_val is the percentage-panning factor for the audio signal; we can find the individual volumes of the channels as follows:

The initial volume: volume = 1

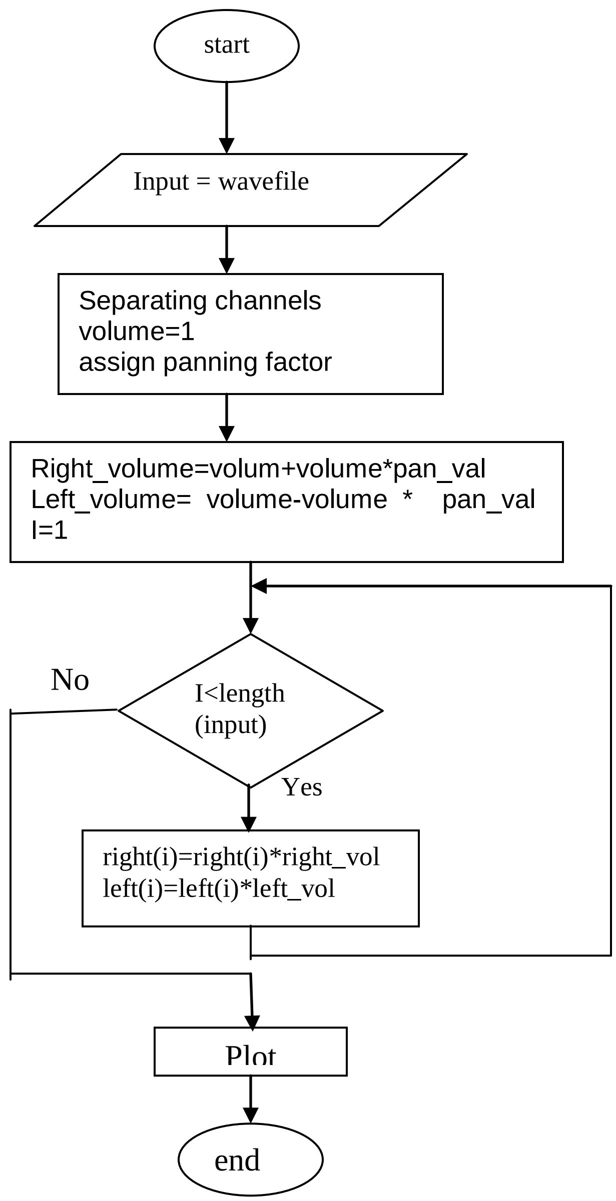
Volumes to be multiplied to the individual channels are given as:

vol_left = volume - volume*pan_val (4.2.1.-a)

vol_right = volume + volume*pan_val (4.2.2-b)

The factor ‘pan_val’, determines the channel with higher amplitude(amplified) and that with the lower amplitude (de-amplified), i.e. if pan_val is equal to zero, both channels have equal volumes. If the pan_val is greater than zero, right channel will be amplified and left will be de-amplified. Similarly pan_val less than zero causes an amplification in left and de-amplification in the right channel (as in equation. 4.2.2-a and b).

Balance



3. Phasing

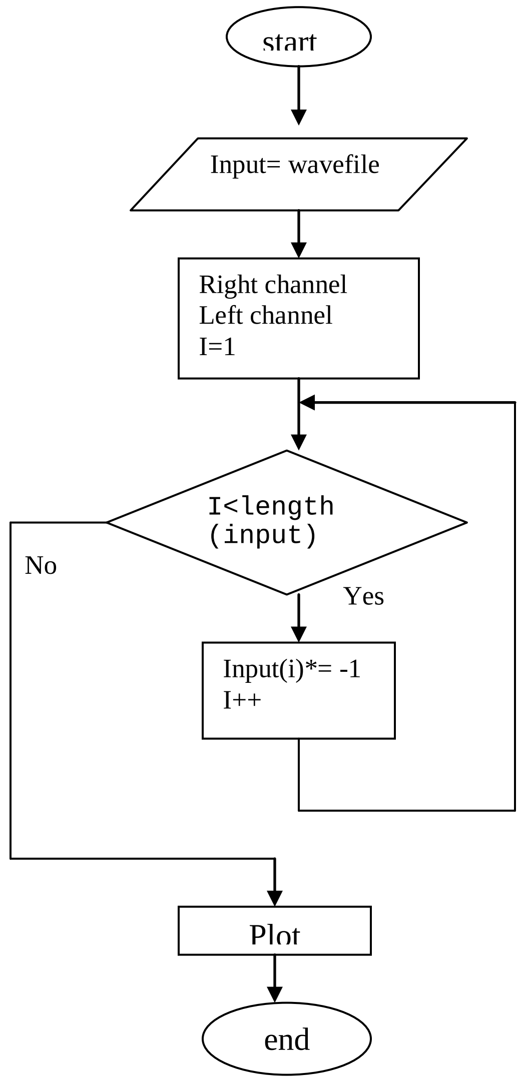
Phasing is another effect that comes under the amplitude category. Phasing is something that carries a key significance in some of the other advanced effects and noise reduction techniques. Phasing actually is inverting the phase of the audio signal, this is helpful while canceling noise from the audio signal, The technique is to find out the frequency of noise in the signal, filter out the noise signal from any one channel, apply phasing to it, and add it to the other channel. Similar is the way for an effect called Karaoke, which removes the vocals from a song.

Phasing is a simple effect and is defined as:

```
Input signal    x
                for i=1 to length(x)
output signal          y(i)=x(i) * -1
                end
```

The idea is simple; as each sample of the input signal is multiplied with -1, in order to invert its phase.

Phasing:


start
Input= wavefile
Right channel
Left channel
I=1
I<length
(input)
No
Yes
Input(i)*= -1
I++
Plot
end

4.2: Delay-based effects

These effects include some of the real-life phenomena, which are sometimes required to be applied to audio, in order to simulate the same, or introduce some kind of an enhancement by applying any one some of them to the audio signal, virtually creating an environment in the audio. These effects hold their importance in the music industry. The effects under this category are as follows:

- Flange
- Echo
- Reverberation
- Chorus

1: Flange

Flange is the simplest of the time delay-based effects. Flange is defined when a sound is reproduced with a very short delay and is mixed with the original sound, giving feeling as if there are two synchronized sounds being played together. Flange is one of the effects, which are most widely used in the music industry, especially pop and trance genres.

Flange can be applied by reproducing the audio signal, delaying it by the delaying factor and adding it to the original one.

Consider an audio signal that has to be flanged, say the delay factor for flange is 100 milliseconds. We convert this time delay into number of samples to be delayed. This can be done using the sampling frequency of the signal. Say it is 8000Hz. This follows:

No. of samples in 1 sec = 8000

No. of samples in 1 msec = 8

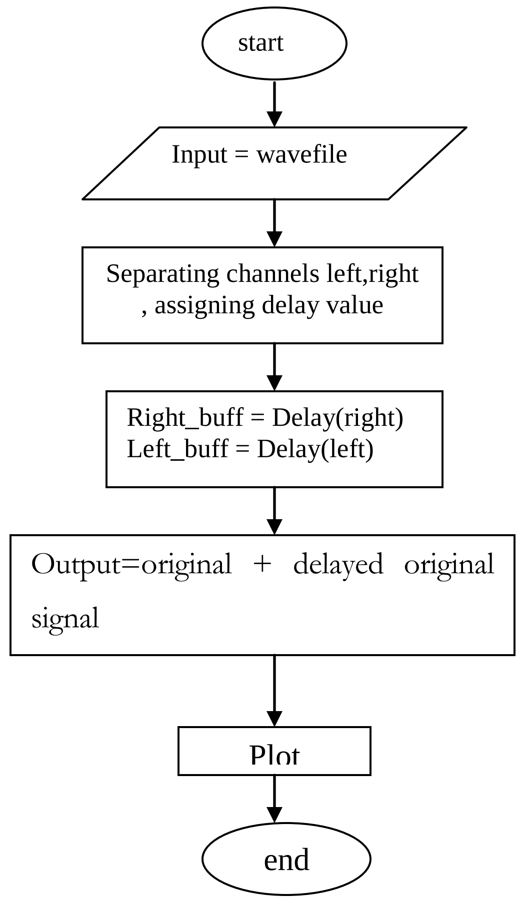
No. of samples to be delayed = 8 * 100

Which gives 800 samples? Thus in order to introduce a delay of 1msec in an audio signal with a sampling frequency of 8000Hz, the signal should be delayed by 800 samples.

Once the signal is delayed through the appropriate number of samples, it is then added to the original signal, in order to get the Flange as follows:

Output = input + delayed_input

Flanging


start
Input = wavefile
Separating channels left,right , assigning delay value
Right_buff = Delay(right)
Left_buff = Delay(left)
Output=original + delayed original signal
Plot
end

2: Echo

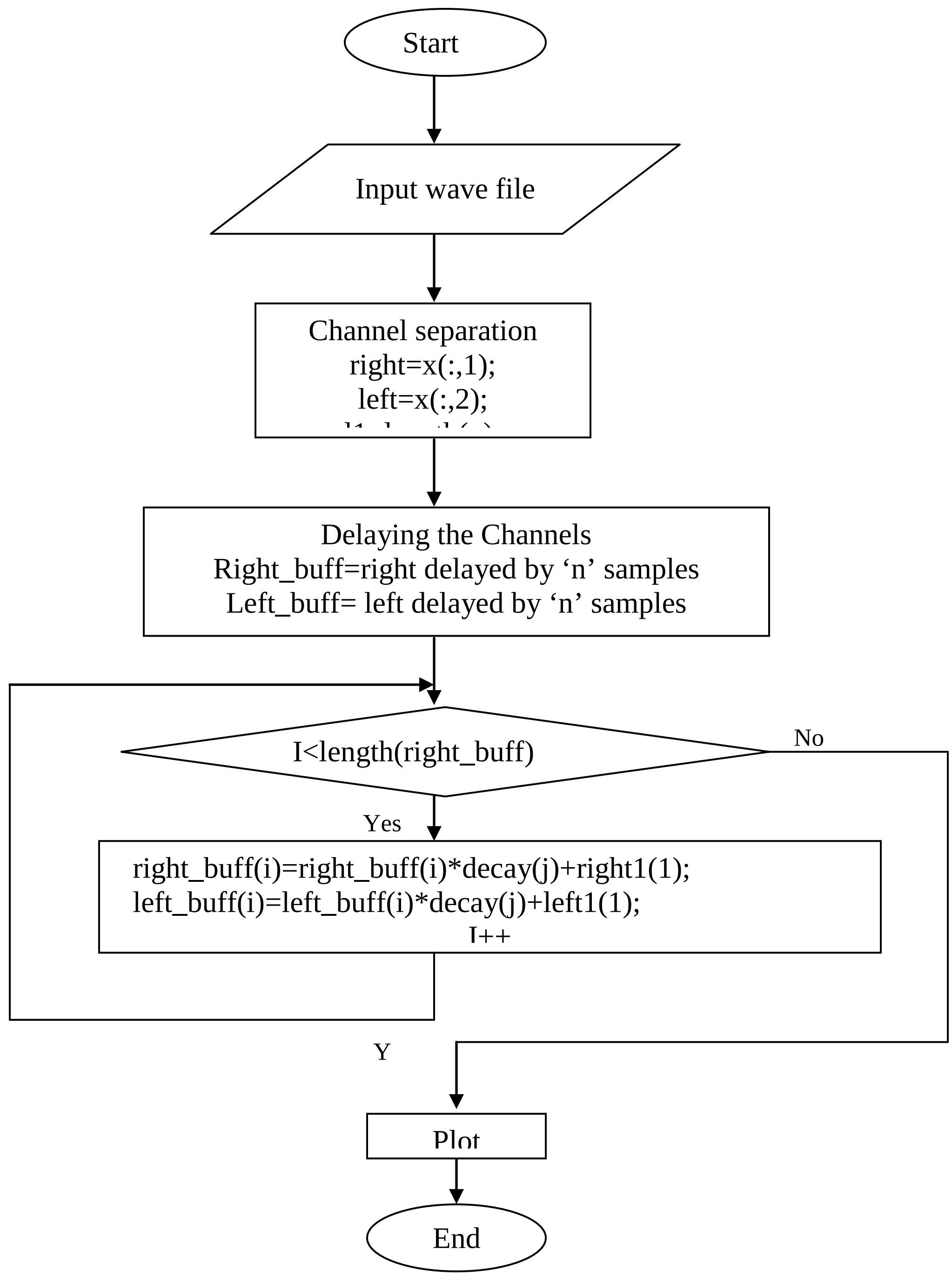
When you roar in a valley, or loudly call someone, you hear your own voice after a short delay once or twice. This reflected voice is what is called the 'Echo'. Echo generally implies a distinct, delayed version of a sound, as the listener would hear with a delay more than one or two-tenths of a second or even more than this. Each echo, or the delayed version of the sound, is easily recognizable by the listener as the delay is quite significant. The reason for these reproduced versions of the sound is that when the sound travels a medium, and strikes hurdles in its way, it either bounces back or reflects at some other angle, and after a delay, reaches the listener again. The delayed version is attenuated as well, because it loses some of its energy on its collision with the hurdles.

The simulation of echo is quite simple, based on the phenomena. It can be done simply by re-generating the audio signal, delaying it by a suitable interval, multiplying it with the decaying factor and adding it to the original one. The output comes out to be a mix of the original sound and its delayed and attenuated versions. This is mathematically defined as follows:

Output = Input + Delayed_input * decay

Here the decay is the percentage energy absorbed by the hurdle from the signal on collision, while the delay is the time taken by the signal to travel from the sound source to reach the hurdle and back to the listener on reflection.

Echo


Start
Input wave file
Channel separation
right=x(:,1);
left=x(:,2);
Delaying the Channels
Right_buff=right delayed by ‘n’ samples
Left_buff= left delayed by ‘n’ samples
I<length(right_buff)
No
Yes
right_buff(i)=right_buff(i)*decay(j)+right1(1);
left_buff(i)=left_buff(i)*decay(j)+left1(1);
I++
Y
Plot
End

3: Reverberation

Reverberation is the result of the many reflections of a sound that occur in a room. From any sound source, there is a direct path that sounds covers to reach our ears. But that's not the only way the sound can reach the ears. Sound waves can also take a slightly longer path by reflecting off a wall or the ceiling, before arriving at ears. A reflected sound wave like this will arrive a little later than the direct sound, since it travels a longer distance, and is generally a little weaker, as the walls and other surfaces in the room will absorb some of the sound energy. Of course, these reflected waves could again bounce off another wall before arriving at the ears, and so on. This series of delayed and attenuated sound waves is what is called reverb, and this is what creates the 'spaciousness' of a room.

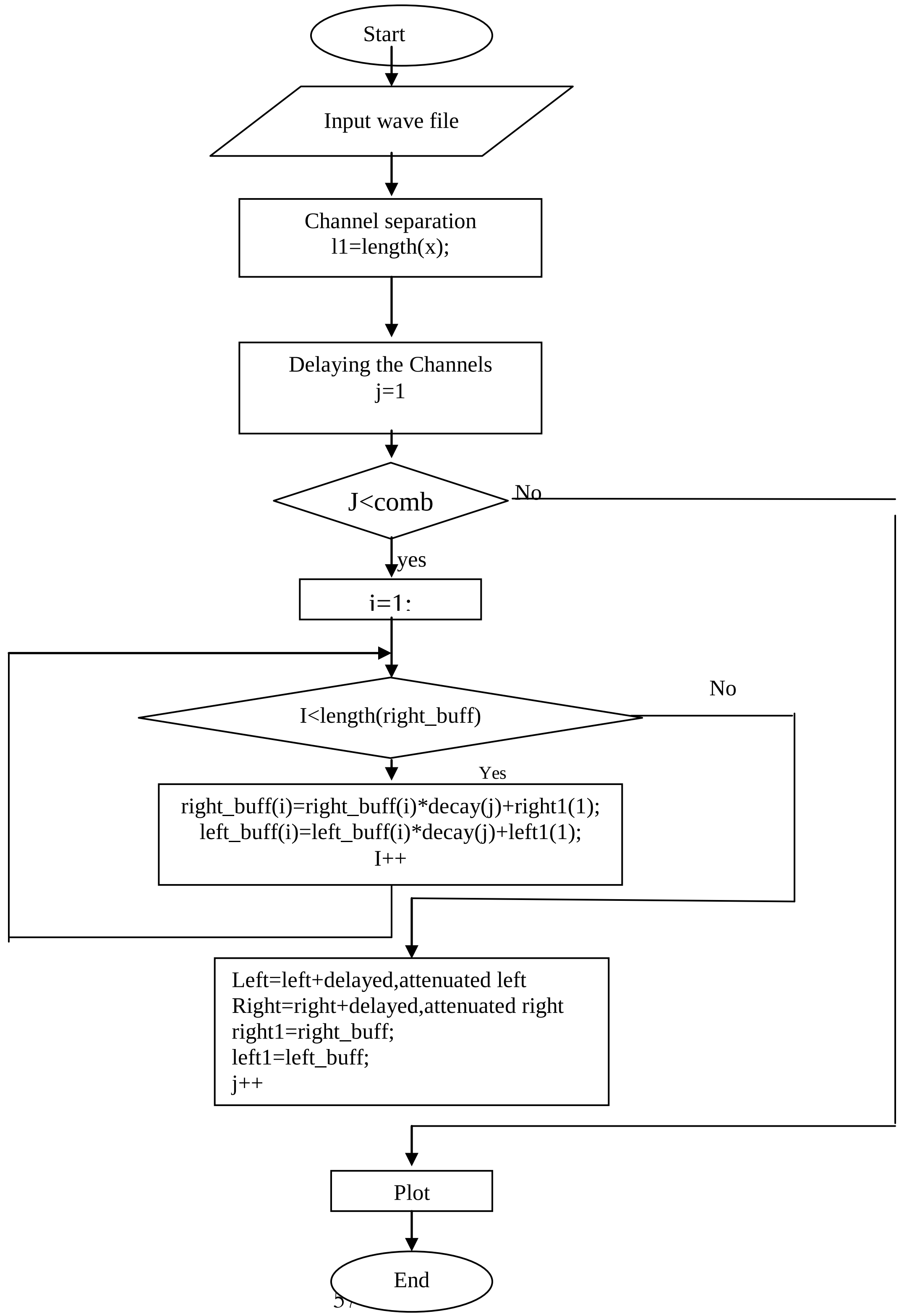
Reverberation can very rightly be termed as multi-echo, the only difference being that this phenomenon is defined for shorter delays, i.e. the delayed versions of the sound are not recognizable separately, as is the case in echo, but one can feel the effect of the sound being repeatedly reproduced.

Consider a room where there is a sound source, and a listener. When the sound is produced it travels in all directions, and reaches the listener ultimately. One path is that it directly reaches the listener. Also, it strikes the walls of the room and reaches the listener with delays, thus causing multiple versions of it. This effect can be simulated as follows:

Output = Input + Delayed_input1*decay1 + Delayed_input2*decay2

The output is the sound that reaches the ears of the listener. Here it is considered that the sound strikes only two walls of the room and reaches the listener. Delay_input1 and Delay_input2 are the delayed versions of the sound, which have been delayed due to different distances of the wall from the listener. *Decay1* and *decay2* are the percentage energies absorbed by wall 1 and wall2, respectively.

Reverb


Start
Input wave file
Channel separation
l1=length(x);
Delaying the Channels
j=1
J<comb
No
yes
i=1:
I<length(right_buff)
No
Yes
right_buff(i)=right_buff(i)*decay(j)+right1(1);
left_buff(i)=left_buff(i)*decay(j)+left1(1);
I++
Left=left+delayed,attenuated left
Right=right+delayed,attenuated right
right1=right_buff;
left1=left_buff;
j++
Plot
End

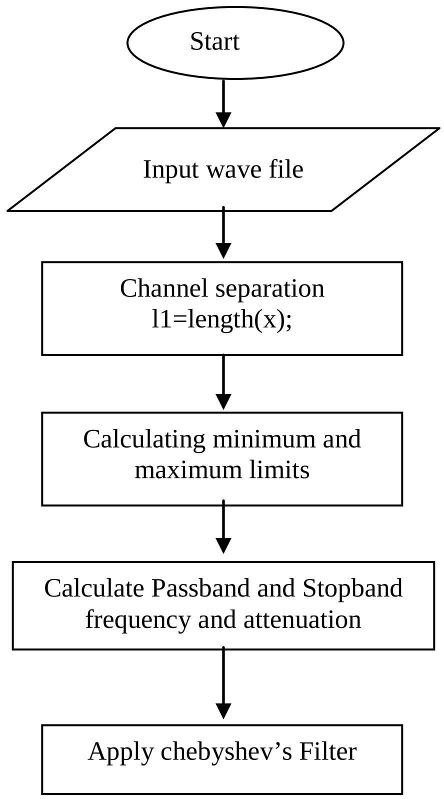
4.3 Quality Augmentation Effects

Equalizer with Chebyshev Order 1 Filter

The Chebyshev Type I Filter is the filter type that results in the sharpest pass band cut off and contains the largest group delay. The most notable feature of this filter is the ripple in the pass band magnitude.

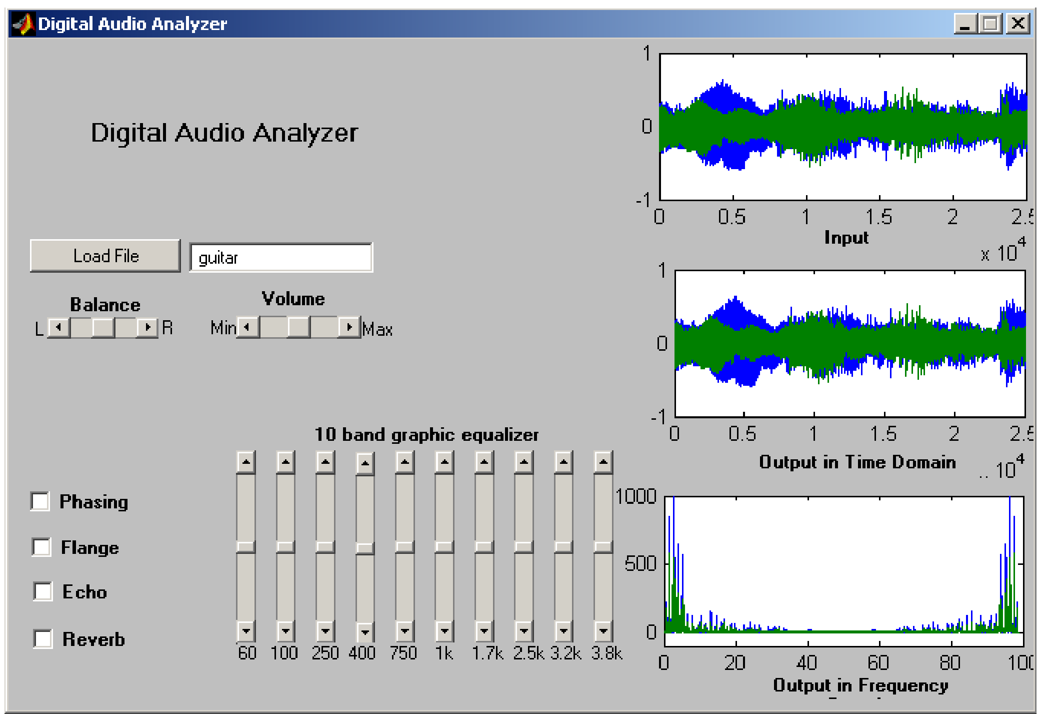
A standard Chebyshev Type I Filter's pass band attenuation is defined to be the same value as the pass band ripple amplitude. However, Filter Solutions allows the user the option of selecting any pass band attenuation in dB's that will define the filters cut off frequency. Filter Solutions also offers the user the option of placing user-defined zeros in the stop band.

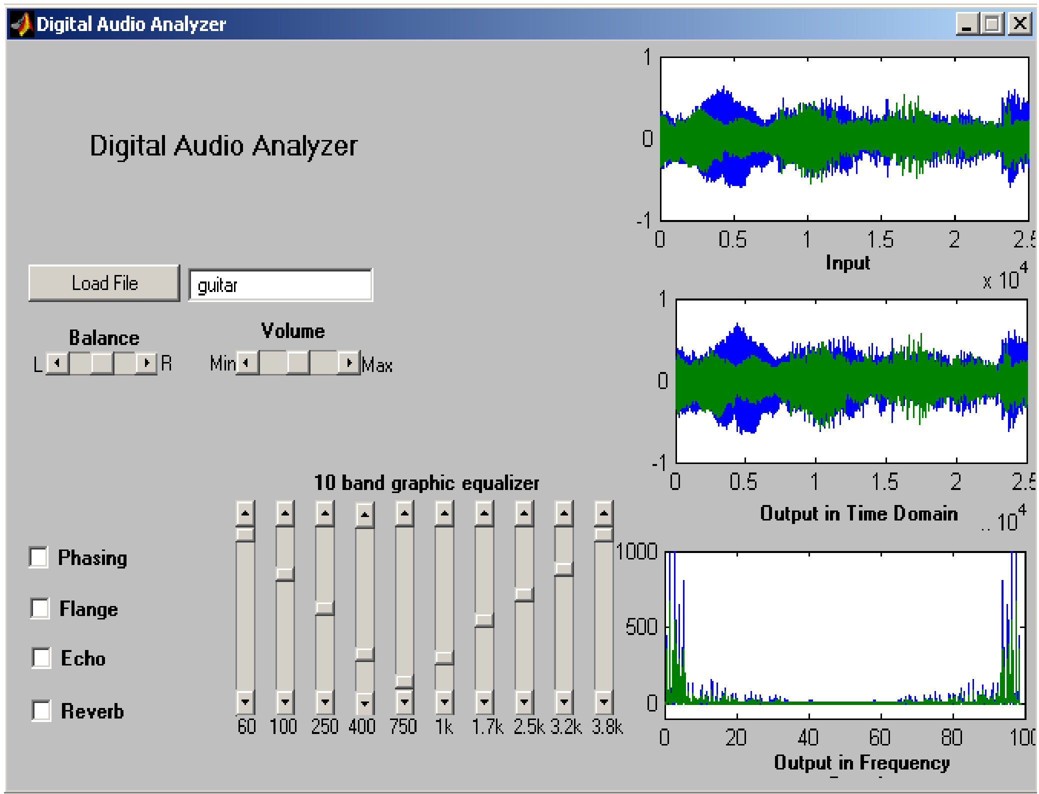
10 band channel Equalizer is developed by using chebyshev order 1 filter. The input signal is either amplified or de-amplified depending upon the placement of defined zeros in the slider bar. A particular range of frequency is selected, minimum and maximum range of frequencies are set. Frequncy is taken as an input from the user, passband and stopband frequencies are calculated and the filter is applied to the calculated frequencies.

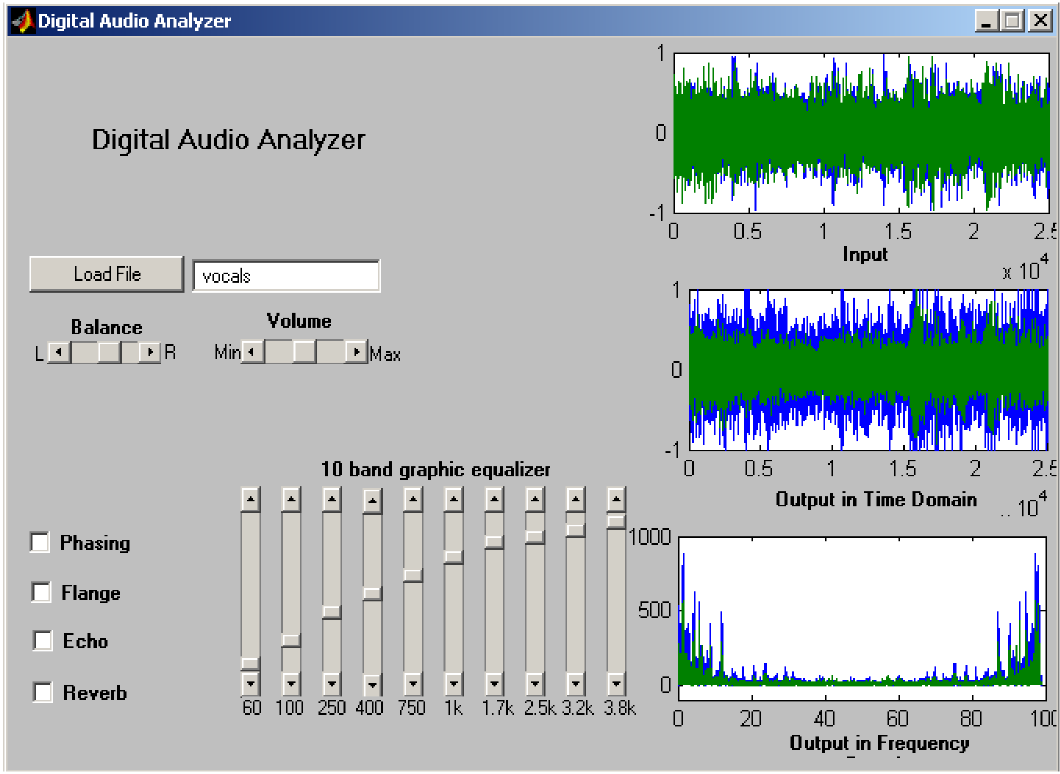

Start
Input wave file
Channel separation
l1=length(x);
Calculating minimum and
maximum limits
Calculate Passband and Stopband
frequency and attenuation
Apply chebyshev's Filter

Chapter 5
Results

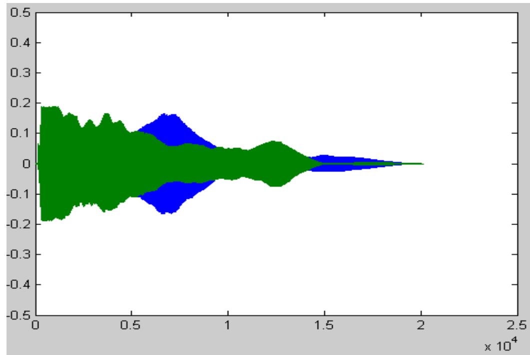
The results of different audio effects are shown with screenshots of the GUI. Different wavefiles are taken as input and effects are applied over them.


Digital Audio Analyzer
Digital Audio Analyzer
Load File
guitar
Balance
L
R
Volume
Min
Max
10 band graphic equalizer
Phasing
Flange
Echo
Reverb
60
100
250
400
750
1k
1.7k
2.5k
3.2k
3.8k
Input
Output in Time Domain
Output in Frequency

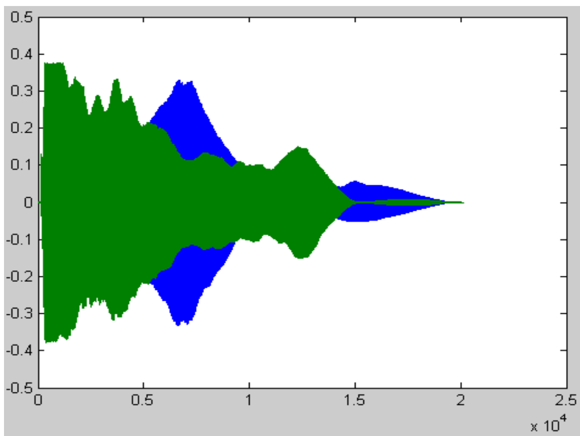

Digital Audio Analyzer
Digital Audio Analyzer
Load File
guitar
Balance
L
R
Volume
Min
Max
10 band graphic equalizer
Phasing
Flange
Echo
Reverb
60
100
250
400
750
1k
1.7k
2.5k
3.2k
3.8k
Input
Output in Time Domain
Output in Frequency



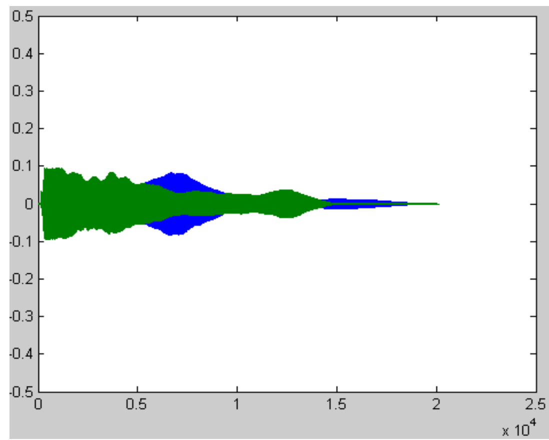
Digital Audio Analyzer
Digital Audio Analyzer
Load File
vocals
Balance
L
R
Volume
Min
Max
10 band graphic equalizer
Phasing
Flange
Echo
Reverb
60
100
250
400
750
1k
1.7k
2.5k
3.2k
3.8k
Input
Output in Time Domain
Output in Frequency



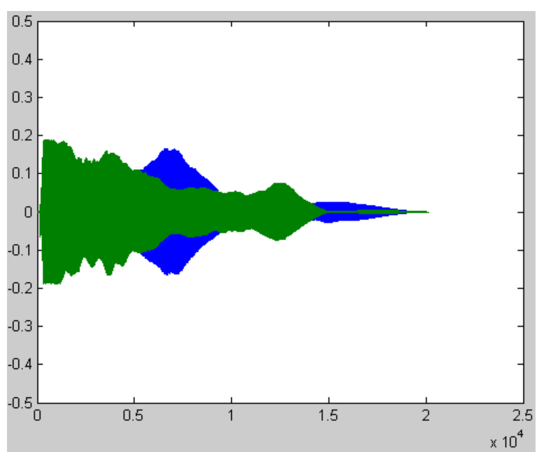
Input



Amplified Output

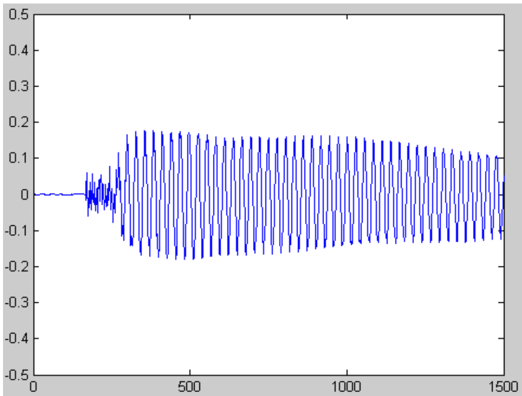


Attenuated Output

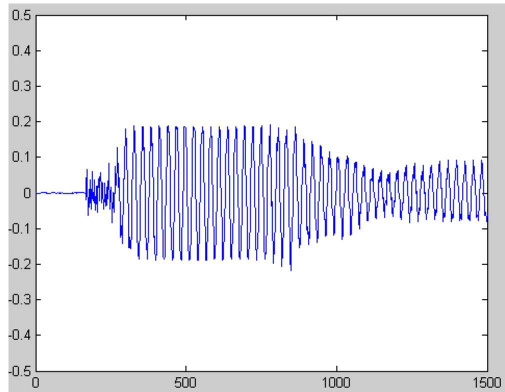


Phased Output

Echo

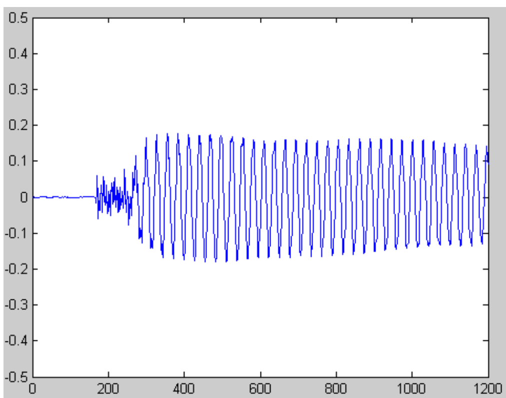


Input

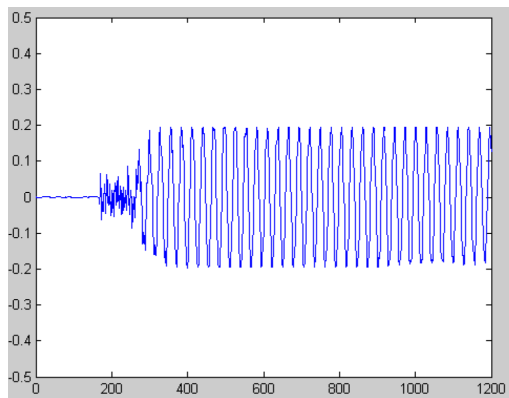


Output

Reverb

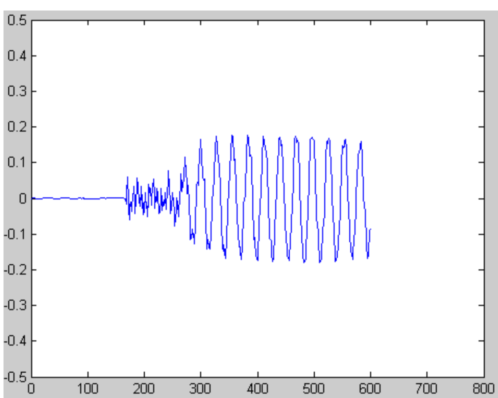


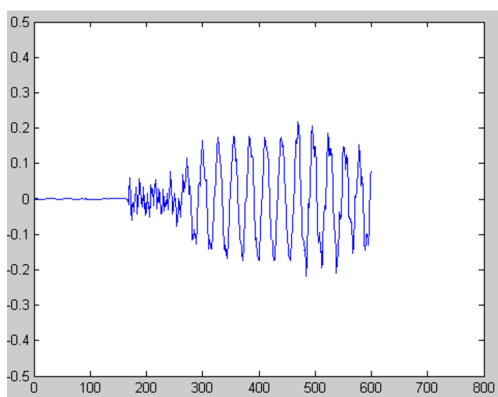
Input



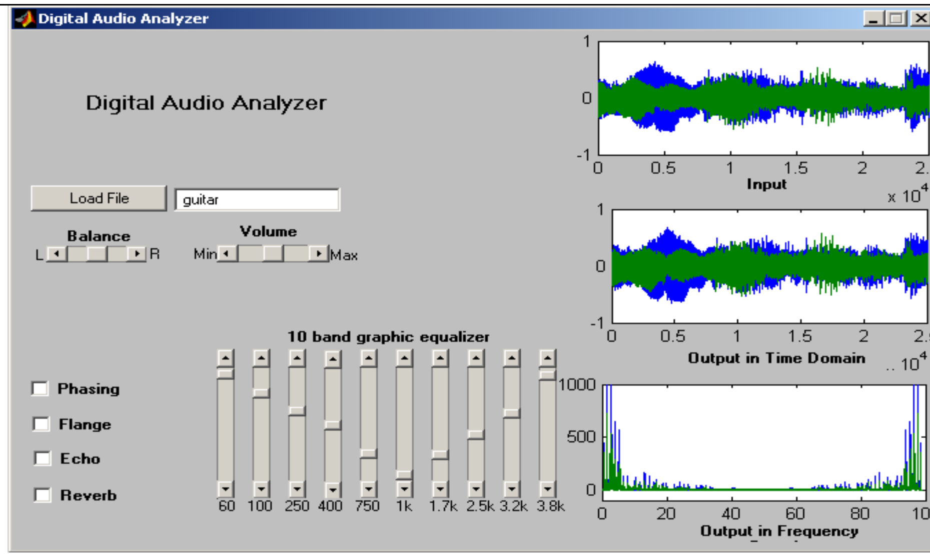
Output

Flange

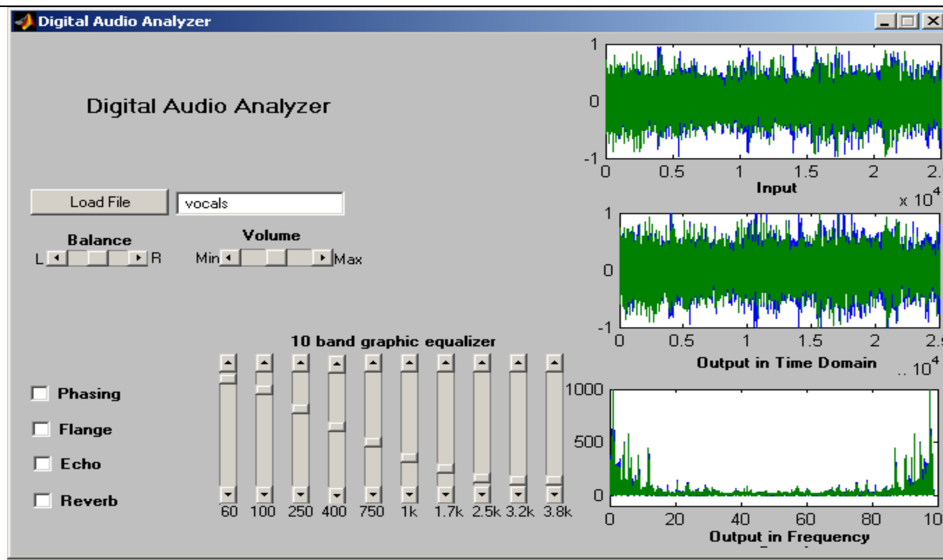




Input Output



frequencies adjusted to de amplify vocals and enhance musical instruments



Frequencies adjusted to enhance bass

Chapter 6

CONCLUSION

The thesis has achieved all of the aims set at the commencement of the project. The design method has been proven to work effectively. Several shortcomings and areas of potential improvement have been foreshadowed through out the thesis.

5.1 Shortcomings and Limitations

Code optimization

The code can become more efficient by doing optimization. Removal of many unwanted noise and other destructive factors will make the effects more effective. As a result the performance of the code will be improved.

Memory usage limitations

Audio files consume much memory to execute. Addition of more effects on those audios can led them to a tremendous usage of memory. Smaller size of RAM will take more time to run any of these files and also sometimes cause data loss or error in the output. As a result, there is a possibility of wrong results. A particular sized audio file could only be used for applying the effects.

5.2 Future Work

It is appropriate to conclude this paper with a brief summary of the future directions this work would take. The project has the capability to adjust more audio effects. The system performs well when small audio files are used but has the capability to work on large files as well, depending upon the memory of the computer. The thesis has proven the design work. It can be implemented in C to speed up the processing time.

Final task of the project is to apply it to the real world. The existing design of the algorithm supports only the floating point for mat. Another task can be to convert these algorithms to fixed point format. Existing system is capable of applying few delay and time based effects. Quality augmentation effects are also included. Karaoke and noise cancellation effects can also be added to enhance the performance of the Digital audio analyzer.

BIBLIOGRAPHY

[1] Doe, John B. *Conceptual Planning: A Guide to a Better Planet*, 3d ed. Reading, MA: SmithJones, 1996.

[2] Smith, Chris. *Theory and the Art of Communications Design.* State of the University Press, 1997.

[3] John G.Proakis & Dimitris G.Manolakis : Third edition Digital Signal processing Principles, Algorithms & Applications .

Sanjit K. Mitra : Digital Signal Processing A computer-based Approach 2nd edition.

Glossary

Frequency :

Infrasound:

Ultrasound:

Decibel:

Compression:

Rarefaction:

Analog (Analogue):

Digital:

Filter:

Equalizer:

IJSBAR
International Journal of Sciences: Basic and Applied Research

IJSBAR
International Journal of Sciences: Basic and Applied Research

International Journal of
Sciences: Basic and Applied Research

Print & Online

Published by:
GSSRR
Global Society of Scientific Research and Researchers
Visit: www.gssrr.org

ISSN 2307-4531 (Print & Online)