Immersive Assisting Tool and Application for Visually Impaired People : Audio Frequency Segmentation in producing different rate of vibrations

by

ZURIATI AZHAN BINTI ZULKIFLI

Dissertation submitted in partial fulfilment of
the requirements for the
Bachelor of Technology (Hons)
(Information Communication Technology)

MAY 2011

Universiti Teknologi PETRONAS
Bandar Seri Iskandar
31750 Tronoh
Perak Darul Ridzuan

CERTIFICATION OF APPROVAL

Immersive Assisting Tool and Application

for Visually Impaired People :

Audio Frequency Segmentation in producing different rate of vibrations

by

ZURIATI AZHAN BINTI ZULKIFLI

Dissertation submitted to the
Computer and Information Sciences Programme
Universiti Teknologi PETRONAS
in partial fulfilment of the requirements for the
BACHELOR OF TECHNOLOGY (Hons)
(INFORMATION COMMUNICATION TECHNOLOGY)

Approved by,

(DR HALABI BIN HASBULLAH)

UNIVERSITI TEKNOLOGI PETRONAS
TRONOH, PERAK
MAY 2011

CERTIFICATION OF ORIGINALITY

This is to certify that I am responsible for the work submitted in this project, that the original work is my own concept as specified in the references and acknowledgements, and that original work contained herein have not been undertaken or done by unspecified sources or persons

Zuriati Azhan Zulkifli

ABSTRACT

This project addresses the problems and limitations that the visually impaired people faced, especially in the area of sports and entertainment. With the advancement of ICT, an immersive assisting tools and application is to be developed to fill in the gap between the impaired people and the society. Other objectives of the project are to develop an assisting tool and application for the visually impaired people so that they can immerse and engage into a specific game of his/her interest, to look into the impact of using audio as the aid in representing the 'real situation', to experiment on audio segmentation into different frequencies and tones as a result of different kinds of vibration as well as to compare and benchmark the studies of this project with related works. The project will be specifically conducted on football game with the motivation to continue the facilities provided for the visually impaired people by the FIFA World Cup 2010, South Africa. The approach proposed is on audio, video and haptics. But this project focuses on audio or sound as it is indeed the best element in assisting the visually impaired people to get immerse and feel just like they are really 'watching' the football game. Assuming the football match will be in the type of playback video to resolve the issues of synchronization and noise of live football match, and it will then being extracted to both audio and video file. The main contribution of this project is to produce the low and high tones and frequencies which will then being used as the input for various types of vibrations by using Smartphone.

ACKNOWLEDGEMENT

First of all, I would like to give my special thanks for those who have in any way made possible for me to complete the Final Year Project.

In particular, I would like to express my highest gratitude to God for giving me the strength, good health in completing this final year project. A highly appreciation expressed to my supervisor, Dr Halabi bin Hasbullah, the Senior Lecturer of Computer Science Department. Also, for the time spent for evaluating and supervising the author all this while. Thank you.

The author is indebted to University Teknologi PETRONAS for giving the author the experience in doing the Final Year Project, being exposed to the presentation, VIVA and SEDEX which exhibition presentation style.

Last but not least to the author's family and friends for their invaluable supports and encouragement to me from the beginning of the project until the end.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1
INTRODUCTION

1.1 Background

With this fast-paced world of technology, there exist lots of tools and applications for the blind or visually impaired people in order for them to compete and capable of doing the same activities as what the normal people do. However, these technologies are not enough as they still having limited access and capabilities, especially in the area of sporting games. As a result, they were isolating themselves and not participating with the society. These will lead to low self- esteem and decrease the level of confidence to succeed and experience the fullest utility or satisfaction in life.

To specifically address the issues of this project, an immersive assisting tools and technology is to be developed which mainly focus on football game. This will allow them to get access to the facilities and experience themselves on the real game thus providing equal opportunity and active participation. A system consists of multimodal elements such as audio, video and haptics are to be combined together in order to resolve the limitations faced by the visually impaired people as well as to achieve the intended objectives.

This research focuses on the element of audio and how does sound contribute in giving the visually impaired people the feel, same as like they are 'watching' the football game. The challenges present in developing the tool include the perception or mental model of the visually impaired people, accessibility, usability and synchronization of the football match for the live game [1]. Assuming the football match will be in the type of playback video to resolve the issues, it wills then being extracted to both audio and video file. The audio file will be analyzed to produce various patterns of tones and frequencies which then will result in different types of vibration by using Smartphone [2].

1.2 Problem Statement

Most of the people are lucky enough to be born with complete sets of 5 senses which are touch, see, hear, feel, and taste. But some of them were born naturally handicapped as a result of fate. Others are having unfortunate life with some forms of disabilities caused by accidents and disease with high cost of surgery and treatment of it [3]. As a result, this group may have segregated their life at an end corner without good understanding and support from the society. Even worst when equal opportunity has not been given to them to do the same as that of normal people, or the opportunity itself has been taken away from them.

When sight is not available, awareness information cannot be obtained through the visual modality. This often limits a visually impaired person's ability to experience all the activities as well as facilities that are available nowadays. In 2010, the world enjoyed the FIFA World Cup 2010 which was held in South Africa. It was the event that everyone was looking forward to see. However, the blind and visually impaired people are having difficulties in enjoying the football game thus isolating themselves and not participating in the society. For example, to watch it lives from TV or stadium requires support from their friends and family. Even as the match goes on, they could not really follow the match and understand what actually happened as their brain processing might be slower than of normal people. They need their space and time in adapting and understanding the match so that that they can follow and enjoy each and every moment during the match.

With this segregation and limitations issues, an immersive assisting tools and application for the visually impaired people in specifically football game, should be developed in order to help them immerse, enjoy and feel the excitement of the football game. This tool will them can used on other settings, such as playing video games at home, watching movies in theatres as well as for education purposes.

1.3 Objectives

The objectives of this project are:

- To develop an assisting tool and application for the visually impaired people so that they can immerse and engage into a specific game of his/her interest.
- To look into the impact of using audio as the aid in representing the 'real situation'.
- To experiment on audio segmentation into different frequencies and tones as a result of different kinds of vibration.
- To compare and benchmark the studies of this project with related works.
- To help in filling the gaps and limitations of visually impaired people in joining the normal people's activities.

1.4 Scope of Study

This research covers only on audio processing apart from the whole system model of the Assisting Tool and Application for the Visually Impaired People project. The project focuses on the extraction of the audio into different frequencies and tones by using the Fast Fourier Transform which then being displayed. The low and high frequency of the audio from the football playback will be compared and matching with the recorded sound of whistle and 'goal' thus will result in different kind of vibrations. Critical review of related works on sound efficiency and segmentation will be conducted to develop such mechanism which can aid the visually impaired people limitations in 'watching' the game. Assuming the audio processing will be in the type of playback(not live) as there are greater noises, efficiency and synchronization problems and the research focused on the visually impaired people who are not originally blind since birth and blind due to illness and accident in the middle of their life as they already have the imaginations on seeing football games.

1.5 Relevancy and Feasibility of the Project

With the given problem statements and objectives of the project, the author believes this project fulfill the social responsibilities in helping the visually impaired people as it helps them in 'feel' and get immerse with the games and activities that they cannot do with the limitations and impairments. The motivation behind this research is to provide an equal chance and opportunity for them to appreciate the football game, by directly join the crowd in the game venue or be connected to a TV set while the game is being broadcast. Hence, with the aid of audio, video and vibrations, the visually impaired people can enjoy and mix with the normal people and capable of doing the same activities just like the normal people regardless of their impairments.

With the defined scope of study, this project is believed to produce the desired result and being able to demonstrate or simulate to the examiners. A well-planned Gantt chart and research activities will be the guidance for this project in order to achieve the desired result given the short time frame to complete this project.

CHAPTER 2
LITERATURE REVIEW

According to the World Health Organization, 314 million people in the world are visually impaired and 45 million are blind [3]. Formally, a person is legally blind if their central vision acuity is 20/200 or less in the better eye, even with corrective lenses; or if they have central vision acuity of more than 20/200 if the peripheral field is restricted to a diameter of 20 degrees or less. Informally, those who, even with corrective lenses, cannot read the biggest letter on an eye chart are considered to be legally blind [4]. There are various reasons or factors for such impairment, whether originally blind since birth or caused by accidents, diseases or other factors. Hence, the vision which based on perception and imaginary for blind people since birth and the one caused by accident or disease is totally different as the second group have at least see and experience the 'real world' as compared to the first one. This factor is essential to be considered in developing tools and application for the blind and visually impaired people.

With the mission to empower persons with visual impairment by providing them with services & opportunities for greater participation, involvement and integration into society as well as to promote prevention of blindness, Malaysian Association for the Blind (MAB) aspires to create equal opportunities for visually impaired persons so as to enable them to enjoy the same quality of life as the sighted [5]. There are various facilities provided for the visually impaired, one is in the area of Sports & Recreation Service. For sport activities, they can enjoy playing Judo, Ten-pin bowling and Ping Pong while for recreational activities; they can join the Tandem riding, Band music, Gymnasium and more [6]. All these facilities are meant for them to enjoy and experience themselves on the game but still within their group in the association. Hence, for them to be able to join the society, they might need help from the advancement of technology so that they can have the additional 'sense' in order to cater their disabilities within the society.

As for Football Game, it is typically played by healthy normal people on the team basis. Handicapped people are always at the losing end which the best they can do to experienced the game by listening to the match via television or radio. For the blind people that caused by accident or disease, this group has already have the memory of the game formation and tactical in mind. Hence, when they are listening to the game, they can imagine how the game flow as compared to the naturally blind people is. From the previous effort made by FIFA World Cup 2010 in South Africa, which by giving the chance for the blind people to come over to the game venue, by directly joining the crowd has given a great experience for the blinds. This is inline with the FIFA slogan, which say that it is Everyone's World Cup. These special provisions make it possible for the blind people to enjoy the live match with the help of headphones. Six stadiums which each have 15 seats equipped with headphones, and trained commentators will report live on the action happening on the pitch. There were also 15 seats allocated to the sighted guides who accompany the visually impaired and specially trained volunteers who provide assistance within the stadiums [7].

A range of adaptive technologies and devices have evolved since the 1960's to enable the visually impaired to deal with variety situations. The most famous one is Braille by using the DAISY software which has always contributed to the improvement of literacy and independence of the blind people. Other technologies included the Laser Cane, Mowatt Sensor, Sonic Guide VA, Smith-Kettlewell Eye Research Institute Project. All of these devices provide feedback to the user through range of tones and fixed intensity vibrations. However, there are few drawbacks for the existing assistive devices including cumbersome hardware, the level of technical expertise required to operate the devices and lack of portability [8]. To cater all the drawbacks, more researches are being conducted in helping the blind people with the aid of new technology. Such technologies use variety of means such as RFID, IrDA, Bluetooth, WIFI, and Haptics which several solutions have been designed and developed in assisting the blind users [9][10][11][14][18].

Some studies proposed different modes of interaction for blind users who use mobile devices, which implies the implementation of entry modes that use tactile or voice commands, and outputs provided through verbal or iconic sounds [22,23].There is increasingly enormous potential to harness mobile devices (cells and PDAs) capabilities for use in assistive technologies or in developing supporting tool. Since mobile phone has become more and more popular and multifunctional, it provides the best way in resolving the issue of mobility, which the user can bring the device and use it everywhere. There are lots of hand phone-based solutions and with the current features with reliable platforms. Hence, it is believed to be the best devices used for this project. However, to achieve such a result accessibility principles should be applied when developing product or service. Accessibility is a general term used to indicate that a product (e.g., device, service, and environment) is accessible to as many people as possible, including those with disabilities. This is an important feature of systems to allow users with different abilities to access or use them [13].

For virtual environment, the sense of sight and that of hearing have been considered as methods to increase immersiveness. However, there also exist a sense of touch, taste and smell in human senses, and among those senses, the sense of touch has recently come out as an important method for increasing the immersiveness in combination with relevant visual and audio. Especially, in games or virtual realities, the sense of touch greatly helps immerse users into such applications [12][2]. This area of research is called Haptics. As for this project, which is to develop an immersive assisting tools and applications for the visually impaired people, specifically focus on football game, it requires multimodal combination of Audio, Video and Haptics. By using the football playback as the video, it will then be added with sound effects and commentators for audio part and from various frequencies that being extracted from audio to produce various patterns of vibrations by using Smartphone [2].

To further zoom in the Audio part, sound was also shown to be an aid when searching for more detailed information. The benefit of using audio in combination with haptic feedback has been highlighted in studies of cross modal icons for mobile devices. From the results in the studies referenced above it can be concluded that information such as the location of objects or situation could be represented and conveyed by auditory cues. For people with visual impairments, sound is an important information channel. The traditional accommodation for visually impaired users to access data is to rely on screen readers to speak the data in tabular forms. While speech can accurately describe information, such data presentation tends to be long and hard to realize complex information. This is particularly true in exploratory data analysis in which users often need to examine the data from different aspects. Sonification, the use of non-speech sound, has shown to help data comprehension. Previous data sonifications focus on data to sound attribute mapping and typically lack support for task-oriented data interaction [14].

When listening to a sound, the blind person is attempting to form some images and imagine every move, every situation, where the character is, from the special effects done in the movie. There are also several studies that relates the sound with emotion, where different kind of sounds may affect the person emotionally as each person have different characteristic traits. The audio modality is practically the only way to gain information. In addition to the blind and visually impaired people, people with normal visual capabilities may also benefit from information in the audio format, especially in places and situations where the visual presentation of information is difficult or impossible; the rate, pitch, and loudness of speech or the characteristics of background music could be adjusted to make the news more understandable and enjoyable for a given type of user, e.g., introverted or extraverted. [15] [16].

Other than that, sound also help in terms of description and captioning. Audio Descriptions (AD) are simply additional narrative tracks that describe the current scene or setting. Audio descriptions began in the early days of radio when all users relied on well-composed descriptions to visualize what they could not see. Today, most people watch television as a replacement for radio. Unfortunately, this development has created a large void for the visually impaired that must rely on audio to extract the entire meaning of an audio-visual [17]. By using the Microsoft Synchronized Accessible Media Interchange (SAMI) software, the AD can be the aid for users who are blind or have limited vision. Same goes for captioning. Other than that, audio also helps in terms of speech synthesizer. This applies to most of the devices as in order for the blind user to use the devices or software, they need something to navigate them on how to use it, either by using sound or something like Braille/Haptic. For example, the thesis wrote by the UTP student, with the title 'Electronic Mail for the Blind' used the screen reader concept with the use of short key accessibility and speech synthesizer in navigating the user to the changed instruction in the electronic mail [18].

Sounds that is audible to the human ear fall in the frequency range of about 20-20,000 Hz, with the highest sensitivity being between 500 and 4,000 Hz. The perception of sound is influenced by how the auditory system encodes and retains acoustic information [19]. Two recent studies have discussed cross-modality frequency matching between audio and tactile stimulation of the hand [20] [21]. They found that the subjects tend to prefer pairs having the same frequency for the auditory and tactile stimuli [22].

Segmentation plays an important role in audio processing applications, such as content-based audio retrieval recognition and classification, and audio database management. Audio segmentation is a process that divides an audio file into its composite sounds [35]. Each segment or clip should consist of a single sound that is acoustically different from other parts of the audio file. Several methods have been developed for audio segmentation. There are two types of segmentation approaches namely, classification-dependent segmentation (CDS) and classification independent segmentation (CIS). CDS methods are problematical because it is difficult to control the performance [23]. CIS approaches can be further separated into time domain and frequency-domain depending upon which audio features they use, or supervised and unsupervised approaches depending on whether the approach requires a training set to learn from prior audio segmentation results.

The Time Frequency (TF) transformation can be classified into two main categories based on signal decomposition approaches and bilinear TF distributions. In decomposition based approach the signal is approximated into small TF functions derived from translating, modulating and scaling a basis function having a definite time and frequency localization [37]. Distributions are two-dimensional energy representations primarily used for visualization purposes and cannot be efficiently used for parameterization of the signal. Few examples of TF decompositions techniques include short-time Fourier transform (STFT), wavelets, and matching pursuit algorithms [24].

Natural Frequency is a sound wave created as a result of a vibrating object. The vibrating object is the source of the disturbance that moves through the medium as the movement of energy creates sound and various frequencies movement creates various tones [40]. Any object that vibrates will create a sound. Fast Fourier Transform can be used to analyze The Fast Fourier Transform (FFT) and the power spectrum is powerful tools for analyzing and measuring signals. To further zoom in to Discrete Fourier

Transform, it is an application that commonly use for modern signal and image processing applications with an efficient computing method. The functions of Fourier Transform are to acquire time-domain signals, measure the frequency content, and describe the basic signal analysis computations and more.

According to [], the adoption of vibrations is increasingly used to enhance the user experience. Most of the applications and studies are using the presence of particular frequency bands and timer to produce vibration. However, this method are not efficient enough as they vibrate too often and even not at the meaningful events, for example in this case, not vibrating during whistle or goal sound, but vibrates during silent time as they use timer and frequency bands. The end users will be annoyed and loose interest in watching the game, especially in the case of visually impaired people.

For sound matching, past research efforts use the Hidden Markov Model (HMM) by using the FFT to match the sound with prepared sample [41]. However, this is more on speech recognition where we use the real-time sound input from the user. In terms of real time sound processing, there are several drawbacks by using FFT as firstly, the volume dependency should be handled as depending on the volume level, the results of FFT may vary plus the noise sounds involved. Besides that, by taking the time segment of the target sound from continuous input is not as easy as there are wide range of sounds thus we need to be carefully in filtering out the noises or the silence period so that we can carefully reserve the sound and not to miss it. The efficiency of the computation of the sound analysis and matching also becomes one of the big issues and it is very time consuming and will be a big problems for real time application. Hence, the studies use the digital signal processor for the sound analyzer to increase the computation efficiency. These problems of FFT are very much the same in discrete signal processing by using playback.

CHAPTER 3
METHODOLOGY

Methodology is an essential element in a project or research as it provides the theoretical analysis of the methods appropriate for the proposed solutions. It also shows the steps or ways in order to complete on certain process. In this project, the author defines the methodology into two parts, system methodology and research methodology. System methodology refers to the solution for the project while research methodology is the ways in gathering information and data for this project or how the research is being conducted in order to come out with the proposed solution.

3.1 Immersive Assisting Tool and Application System Model

Figure 1 illustrates the complete proposed solution for the Immersive Assisting Tool and Application for the Visually Impaired people. There will be user interface which connect user to the system. It deals with video part which proposed on Braille description. System control will control the user interface and video playback, which will then being extracted to audio and video. As mentioned earlier, video deals with the user interface which handle the user navigation on the playback. The user can play, stop and repeat the playback whenever they want to. Braille technology will be applied on the screen of the hand phone to ease the user in using the tool.

For the audio part, the sound of the playback will be analyzed by using the Fast Fourier Transform (FFT) to frequencies (as well as other properties) [36]. The frequency of the sound will then is being match to the stored sound in order produce vibration. Generally, most of the applications which use haptic and vibration in order to get the user immerse with the system will use monotonous kind of vibrations. The drawback from this is that, the user, especially the blinds will never know the intense and the real situation that happen if the tool keeps on vibration at same level. There will be no difference between goals or offsite or trial shots as each of this event happen, it will vibrate in monotonous vibration. Hence, the proposed system will overcome this problem by generating different kind of vibrations which take the input from different matching frequencies. All these functions will be embedded in hand phones and using headphone to hear the sound.

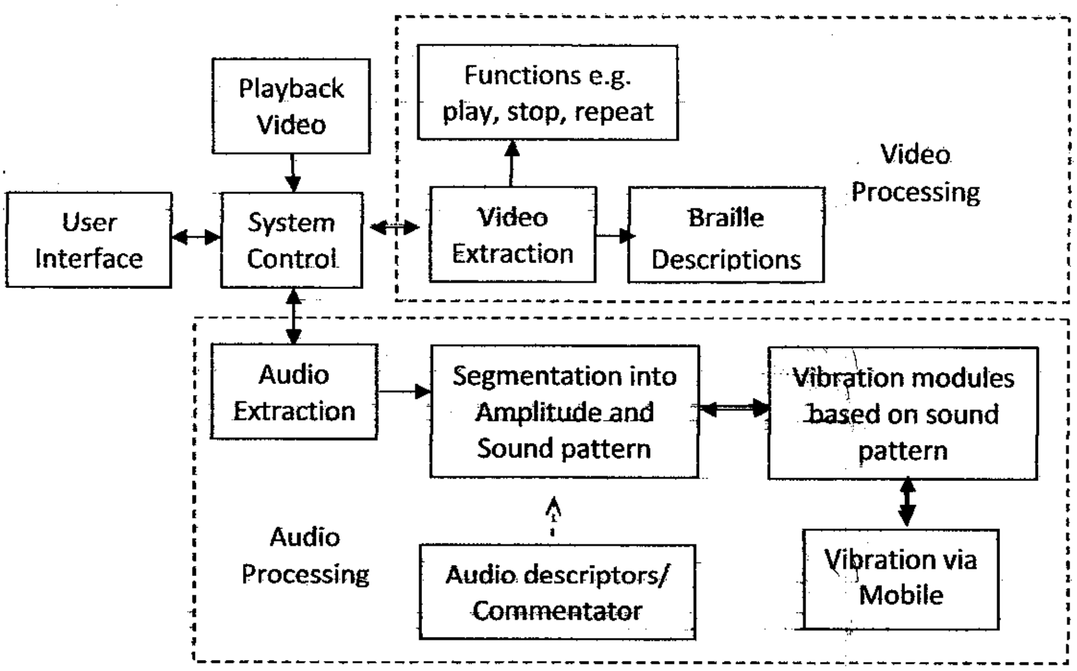


Figure 1 : The proposed methodology for the system

3.2 Audio System Model

The main contribution from this project is to analyze and perform the audio extraction and frequency segmentation which the further being used to produce different kind of vibrations, depending on the football playback. As for football game, there are lots of exciting and surprise moment, for example when it comes to goal, injury, corner kick and so on. The aid of sound from whistle and 'goal' help the blind prople to know the status of the game. From the audio format of the playback (after the audio extraction), the audio then will go through the Fast Fourier Transform (FFT) in order to display the low and high frequency of the playback.

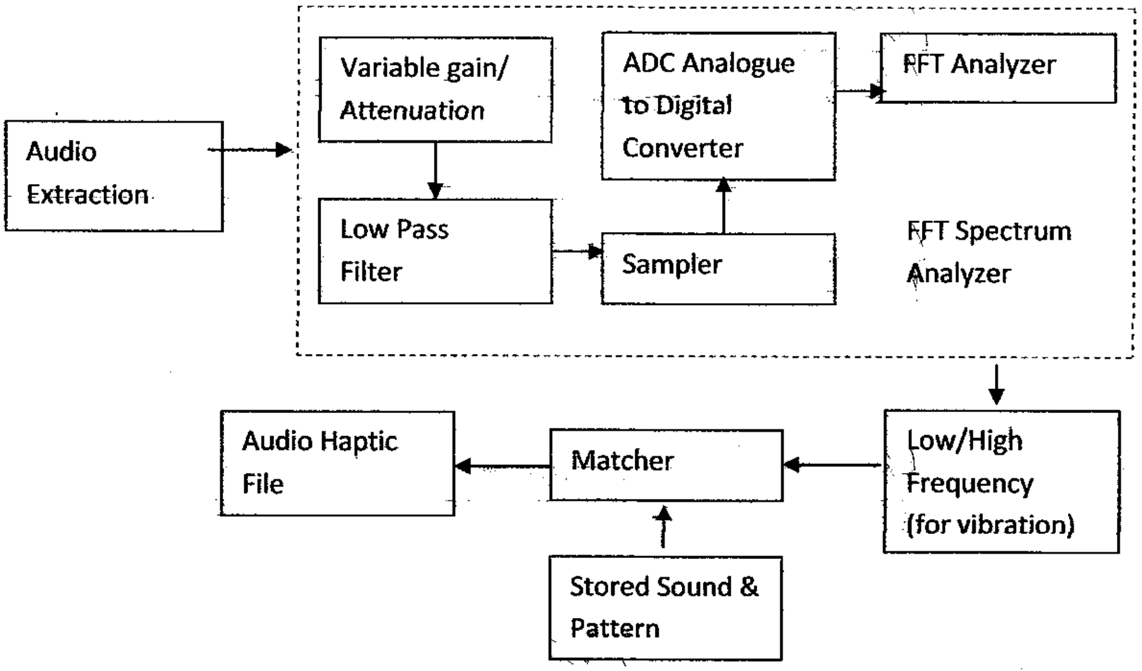


Figure 2: The Audio System Model

There are several processes involved which are first, the variable gain/attenuation to ensure the signal at the right level for the analog to digital conversion. Next, is the Analogue low pass anti-aliasing filter where the signal is passed through an anti-aliasing filter. This is required because the rate at which points are taken by the sampling system within the FFT spectrum analyzer is particularly important. The waveform must be sampled at a sufficiently high rate. After that, the process goes to sampling and analogue to digital conversion. With the data from the sampler, which is in the time domain, this is then converted into the frequency domain by the FFT analyzer. This is then able to further process the data using digital signal processing techniques to analyze the data in the format required. The process will end by presenting the information of display with variety of ways. Hence, all these processes will be done by using the MATLAB software by using certain programming codes.

The author needs to be more familiar with the MATLAB software and perform the audio analysis with the help of Audacity and Sound Forge. This is to be able to view the frequencies of the sound clearer and proceed with the matching process algorithm by using MATLAB. After getting the result of frequencies from the FFT, for example, the frequency of whistle and 'goal', it will then being match with the stored sound pattern of the whistle and 'goal'. The stored sound pattern is where the range of frequencies and the properties of it is being stored and define earlier (the whistle and 'goal' sound). As the playback is running and being analyzed by the FFT to produce the frequencies, it will then being match with the stored sound in order to produce the vibration through the hand phone. By using one of the matching methods as define in the literature review, the process will then complete by producing the audio-haptic file. Figure 3 and 4 shows the example of comparison between two different signals and the example of time-frequency (TF) functions.

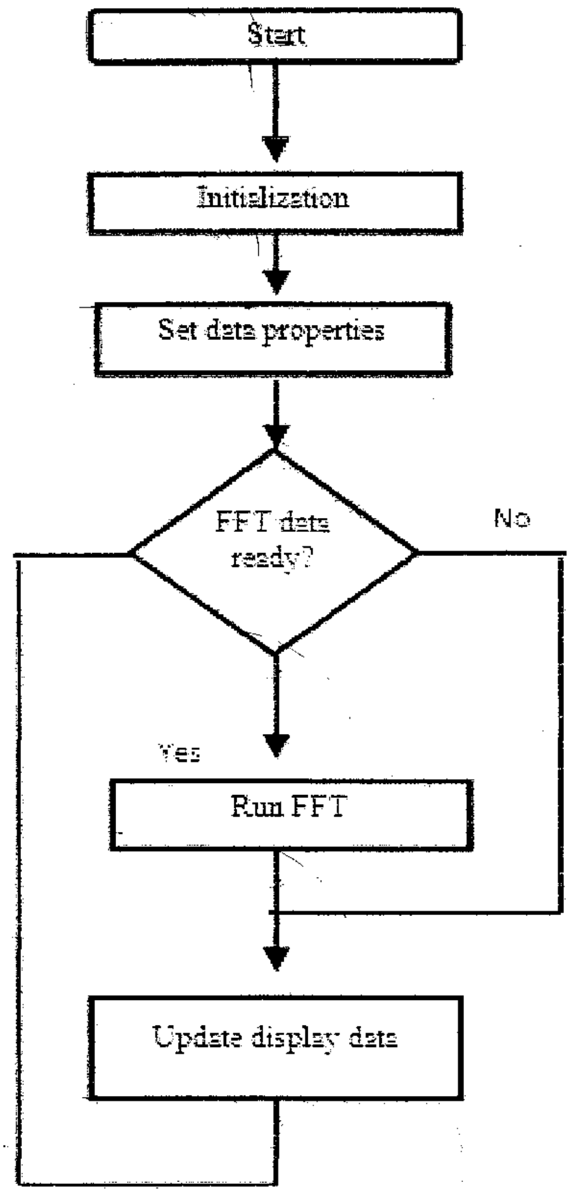


Figure 3: Flowchart for FFT Frequency analysis coding/algorithm

3.3 Research and Data Gathering

Data collection and information gathering is really crucial fro every projects and research. From the data, the author can obtain information to keep on record as well as to make decisions about important issues on the topics. Hence, for this project, the author uses interview, online database, search engine and the websites as well as by referring to books in the Information Resource Centre (IRC).

Interview

An interview is a conversation between two people where questions are asked by the interviewer to obtain information from the interviewee. Interviewing method is the main method use in this project, due to the fact that it is easier to meet with involved persons directly in order to gather for the information. The data captured might be more accurate and meaningful as the movements and the gestures of the interviewees could be physically seen and interpreted. This physical contact might initiate another idea which can lead into another important information acquirement. The team went to Malaysian Association for the Blind to meet Mr Moses. The interview, which took place in MAB for about 4 hours, the team spends time by asking the questions regarding each of the individual projects. Due to time constraint, not all questions are being answered as the team has a lot to ask to Mr Moses. However, the team was amazed by looking the way Mr Moses use all the technologies, and how well he move and walk around. He agreed with all the projects and proposed for few recommendations.

- **Research by using online databases**

Most of the researches and students are using online database or E-Resources to access for information and data gathering. It is an online database that the university have subscribed. It is indeed the best source in getting information as it provides lots of journal papers and research appaers which being wrote by all the researches around the world. In fact, all the papper projects are being revised by the board of committee few times before released it on the database. Hence, the author used the online database as the primary source in doing research ŋ gathering the infromation. Among the online databases are, IEEE, ACM, Emirald, Science Direct, Springerlink and more.Figure 8 shows the chart on frequenly used E-resources in doing research.

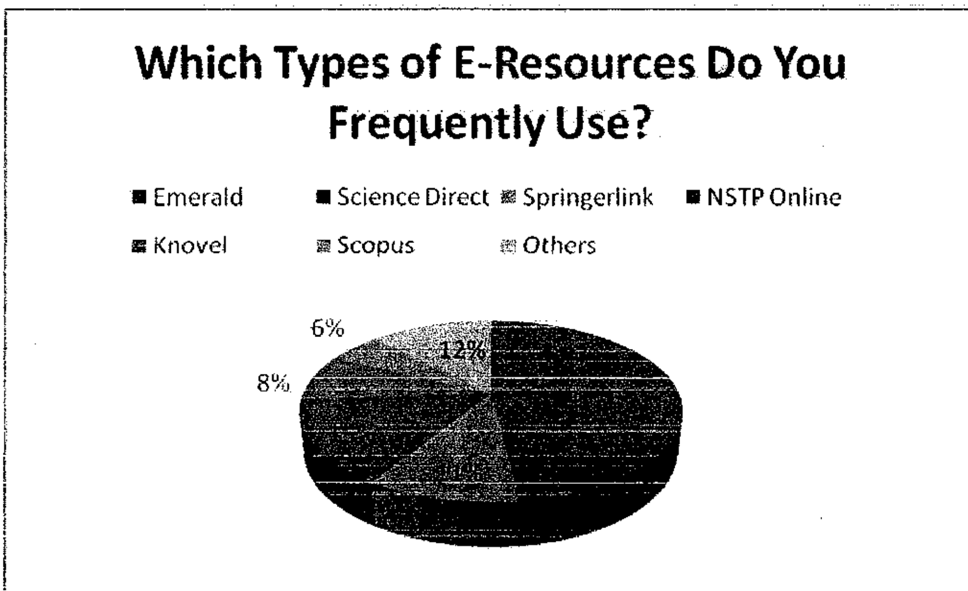


Figure 4: Respondents view on types of E-Resources they frequently use

- **Online Journals and Websites, Search Engine**

  Other than that, the author also uses the online journal, websites as well as search engine in gathering the data. Besides using the search engine like Google, Yahoo, MSNSearch and Altavista, the author also use to refer on the books in the UTP Information Resource Centre.

3.4 Project Activities

Phase 1: Analysis and Design (Critical Review of Related Works)

The project is initiated with a detailed background study for the whole system of immersive assisting tools and application for the blinds. Then, related works on the whole system and specifically on the audio part is being conducted. Its definition, the use of the best tools and concept to be used is reviewed; the use of Fast Fourier Transform, the matching part. Design the proposed solution (methodology) on the big system as well as the audio part.

Phase 2: Simulation Development

Preparation for the audio setup includes identifying and planning on the feasibility and procurement of tools. Start to execute the process by using the Fast Fourier Transform and see the results (the frequency of the audio file). Modeling and analysis tools will be used to improve the viability of the proposed algorithm. Store and capture the frequency of certain sound to be used in the system. At first, use the sound of whistle and 'goal' sound. This phase completes when the audio frequencies match with the stored frequency/tone.

Phase 3: Test- bed development

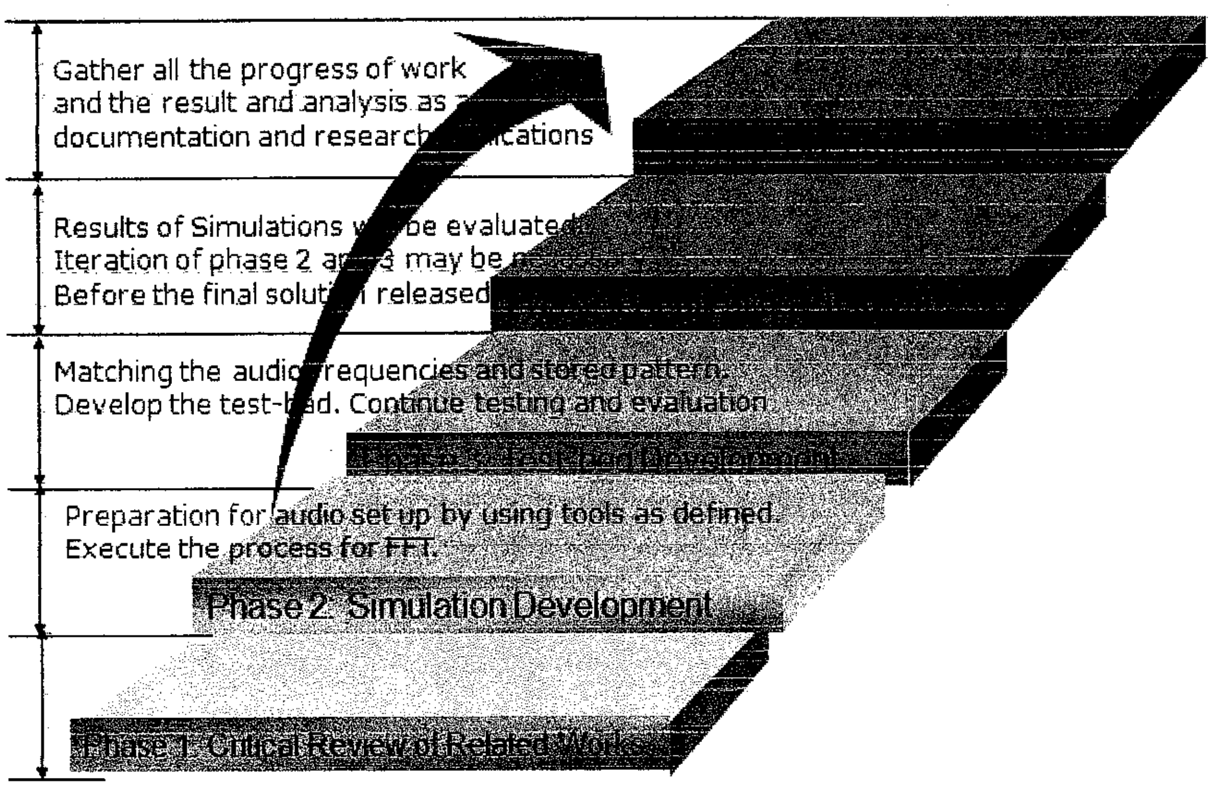
In this phase, test by matching the frequency of the audio file and stored pattern. Use the audio and match fundamentals. Observe and think of how to correct or add the efficiency of the system. Develop the test-bad of the assisting tool and its application, which is specifically on audio part and continue on testing and evaluation.

Phase 4: Application prototype development

Results of the simulations will be evaluated as discussed in related works. Iterations of phase 2 and 3 may be necessary to introduce fixes in the optimization mechanism before the final solution can be released. Continuous integration and enhancements are applied in phase four. Then only the development of the system prototype starts and demo of the system prototype.

Phase 5: Documentation and research publication

Gather all the progress of work and the result and analysis as a documentation and research publications.



3.5 Key Milestone & Gantt Chart

See appendix

3.6 Tools & Equipments

-Fast Fourier Transform Spectrum Analyzer

-Oscillators

-Audio Extractor

-MPEG/MIDI

-Audacity

-Sound Forge

See appendix

CHAPTER 4

RESULTS, ANALYSIS & DISCUSSION

4.1 Analysis & Results

Audio Spectrum Analysis 1: FFT by using MATLAB software

Fast Fourier Transform (FFT) can be done by using MATLAB software by using certain algorithm and coding. It can perform the spectral analysis by using sampled data and view it with other properties such as amplitude, frequency, time, power and more. The FFT allows the audio file to be efficiently being estimated into the frequencies component from a discrete set of values sampled at a fixed rate. The spectral analysis has been done to the football playback which consists of goal sound and whistle sound. The purpose of analyzing the audio file is to get to know the properties of the audio file such as amplitude, frequency, power and time which the value can be used as input to perform vibrations. The three audio files are as below:

- FYP_audio_0001.wav (the full football playback which consist of the goal and whistle sound)
- FYP_goal.wav (the separate sound of goal)
- FYP_whistle.wav (the separate sound of whistle, extracted from the full playback)

These three audio files has been cut and edited by using Windows Movie Maker as sound of goal and whistle itself comes from the full playback. Hence, performing FFT codes by using MATLAB to the three audio files has come out with the graphs. The time scale in the data is compressed by a factor of 10 to make the audio files more clearly audible. The following reads, plots, and plays the data:

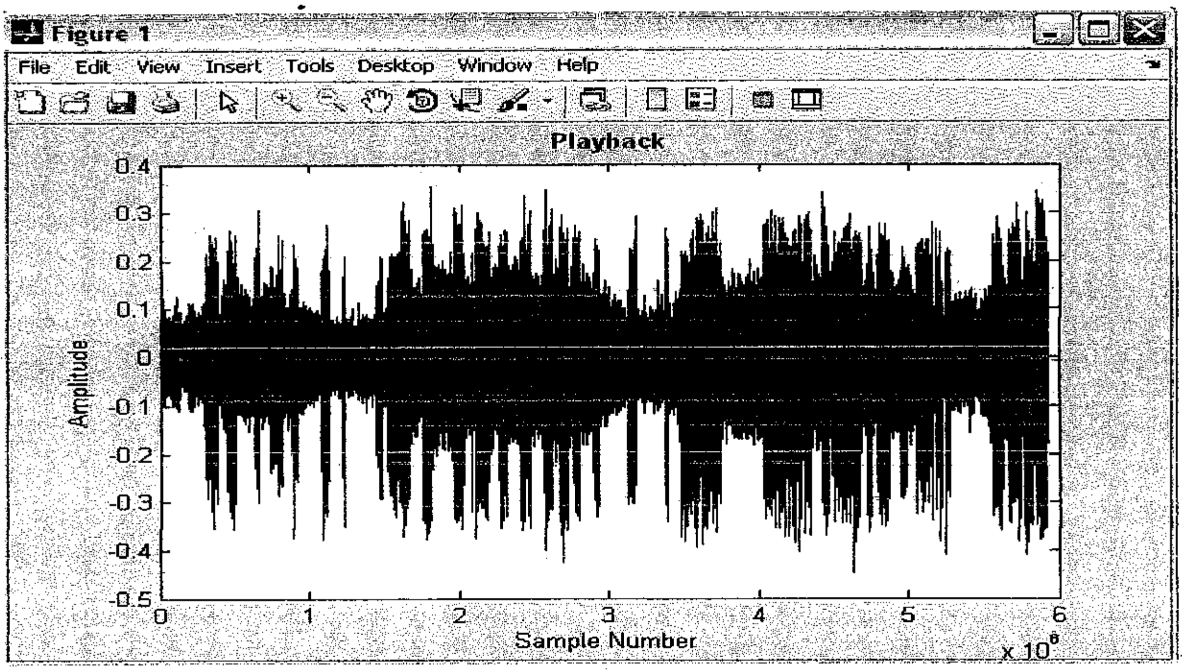


Figure 4: The analysis of full playback

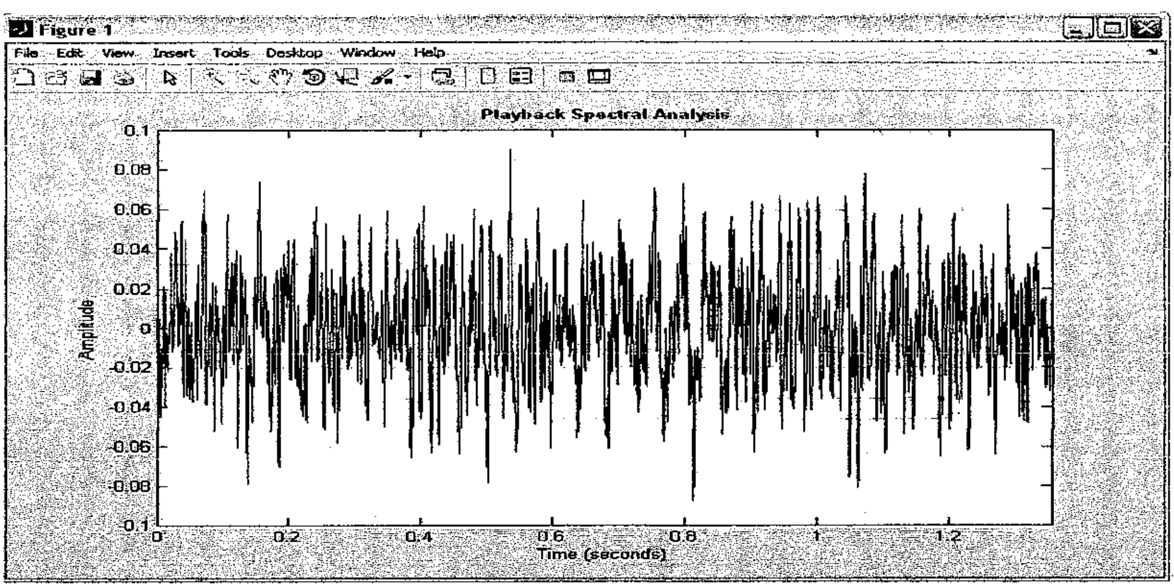


Figure 5: The analysis of the first portion of the playback to see the graph clearer

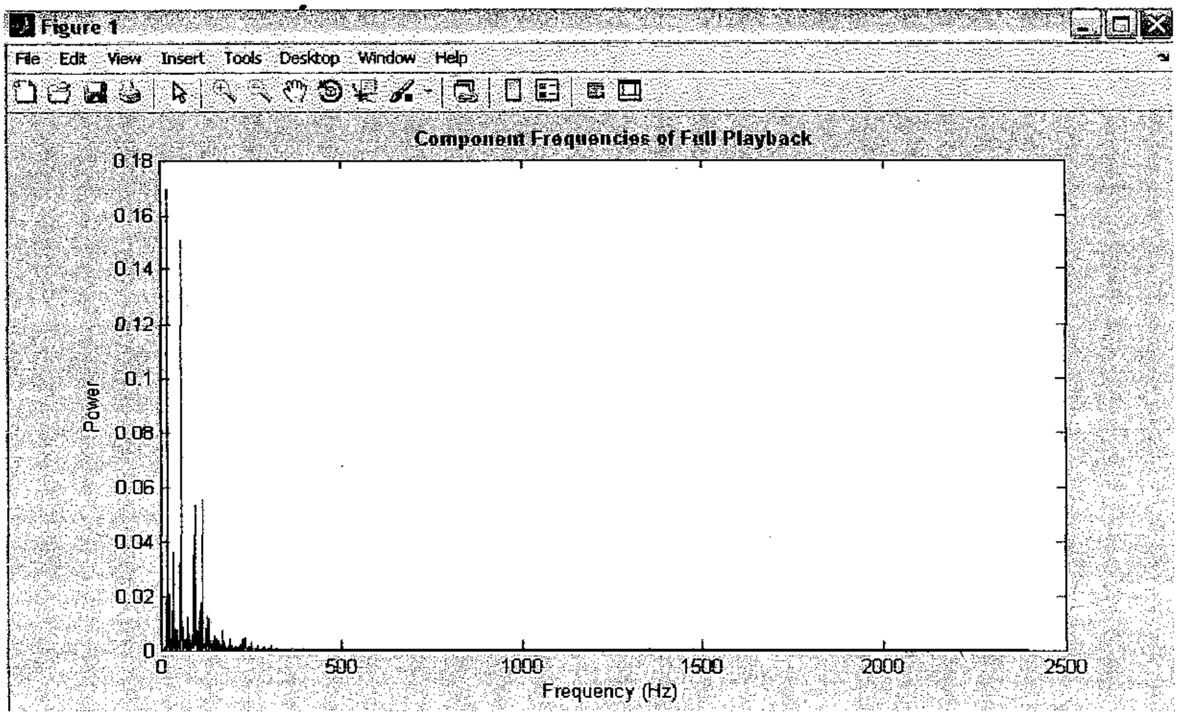


Figure 6: The FFT of the football playback

From the audio spectral analysis above, the FFT does not directly give the spectrum of a signal as it can vary dramatically depending on the number of points (N) of the FFT as well as the number of periods of the signal that are represented. There are common problems when performing the FFT by using MATLAB with regards to the Nyquist frequency. According to the Nyquist frequency, half of the sampling frequency of a discrete signal processing system [2]. The aliasing problem can be avoided by having the Nyquist frequency greater than the bandwidth or maximum component frequency of the signal being sampled. Other than that, as the FFT consist of information between 0 and fs (sampling frequency), the sampling frequency must be at least twice to the highest frequency components. This is called, the Nyquist rate where the signal spectrum should be symmetrical for both positive and negative frequencies. [3].(can be seen from the codes where the spectrum from –fs/2 to fs/2).

The same process has been done to the other two audio files which are the whistle and goal sound. The frequencies of the sound, obtained from the spectral analysis can be use for the matching part or algorithm to perform the vibrations. Below are the analysis of amplitude, frequency, sample number and time for both audio files.

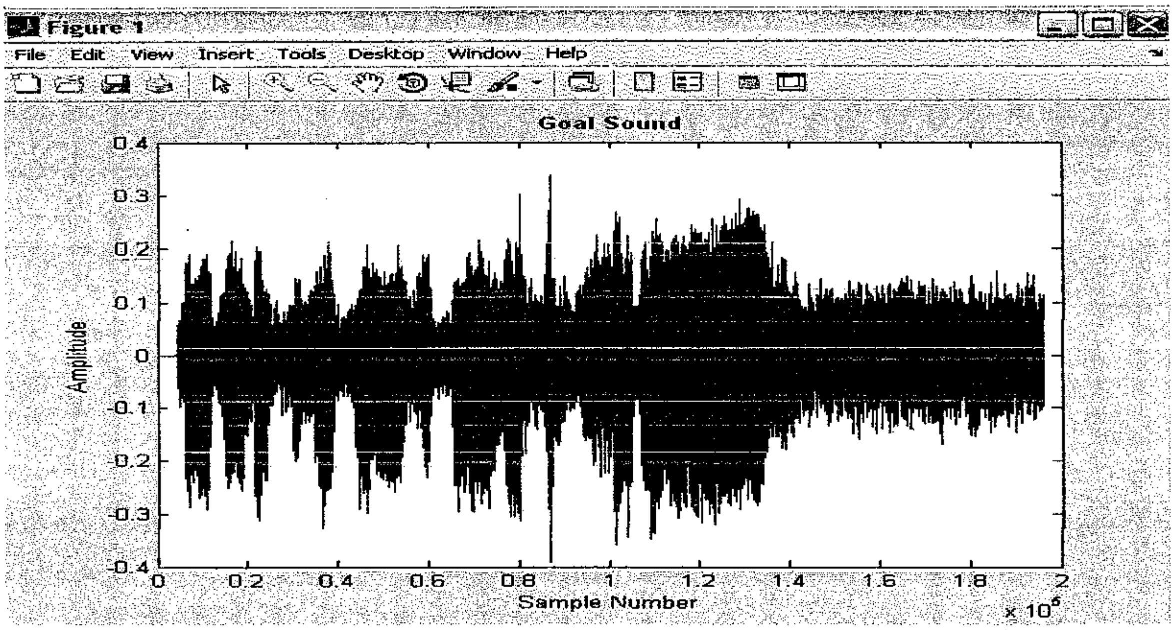


Figure 7: The amplitude of the goal sound

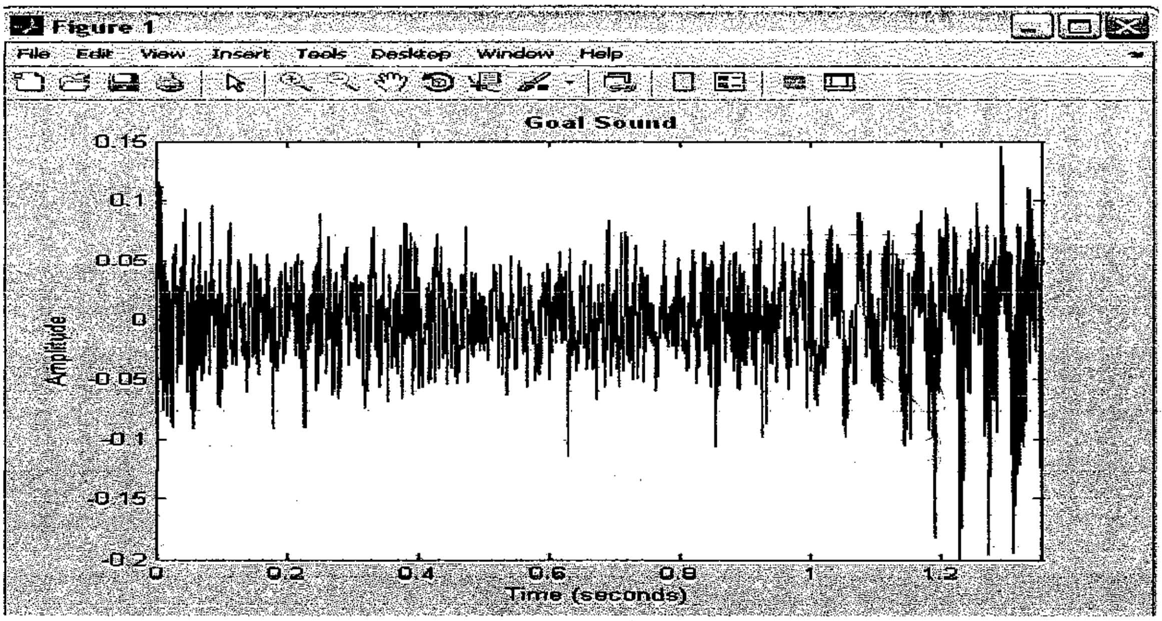


Figure 8: The analysis of amplitude over time

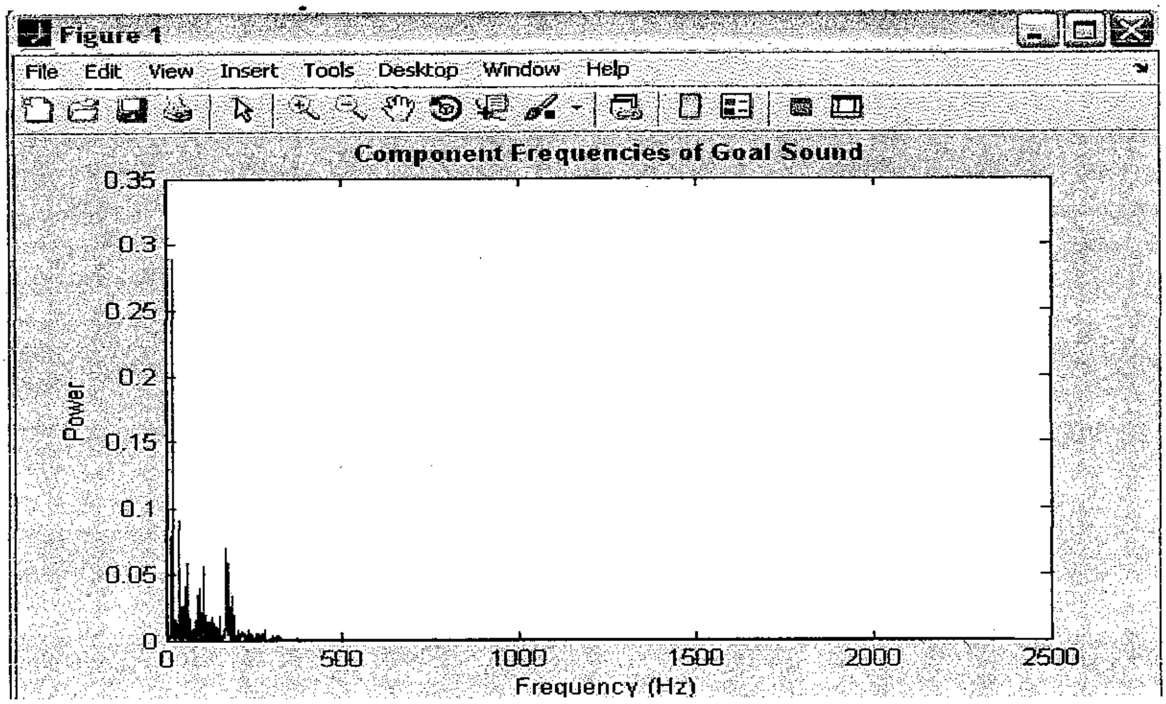


Figure 9: The FFT of the goal sound

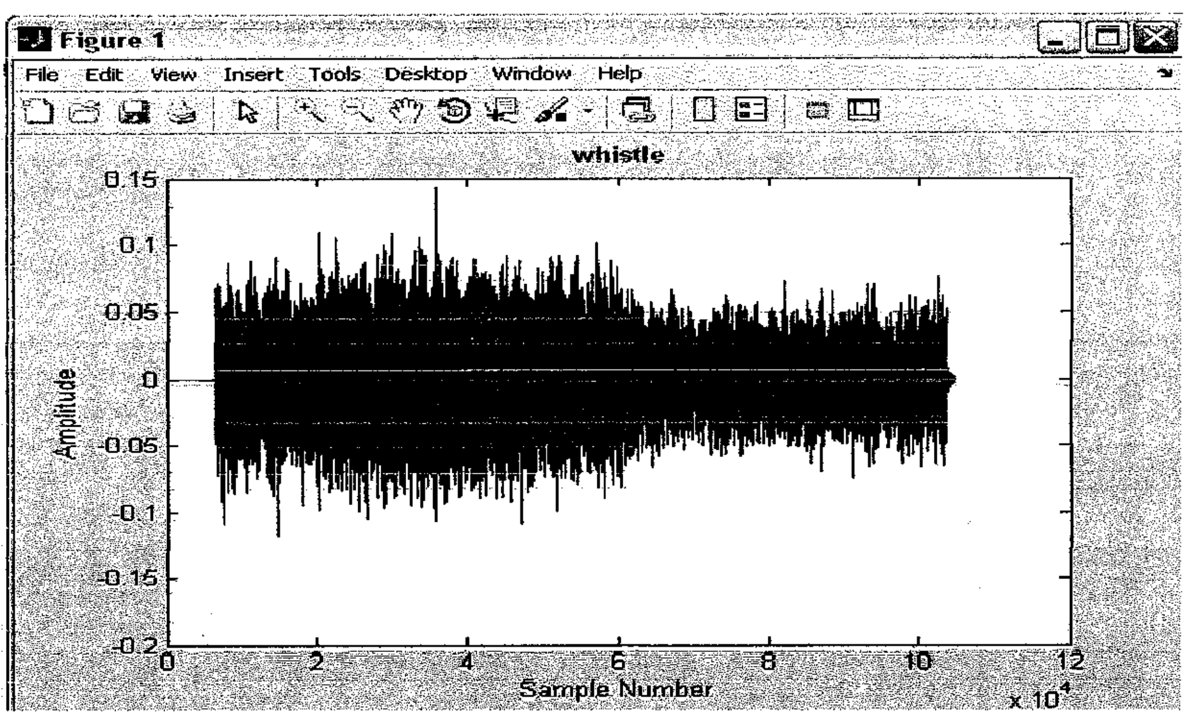


Figure 10: The analysis for the whistle sound

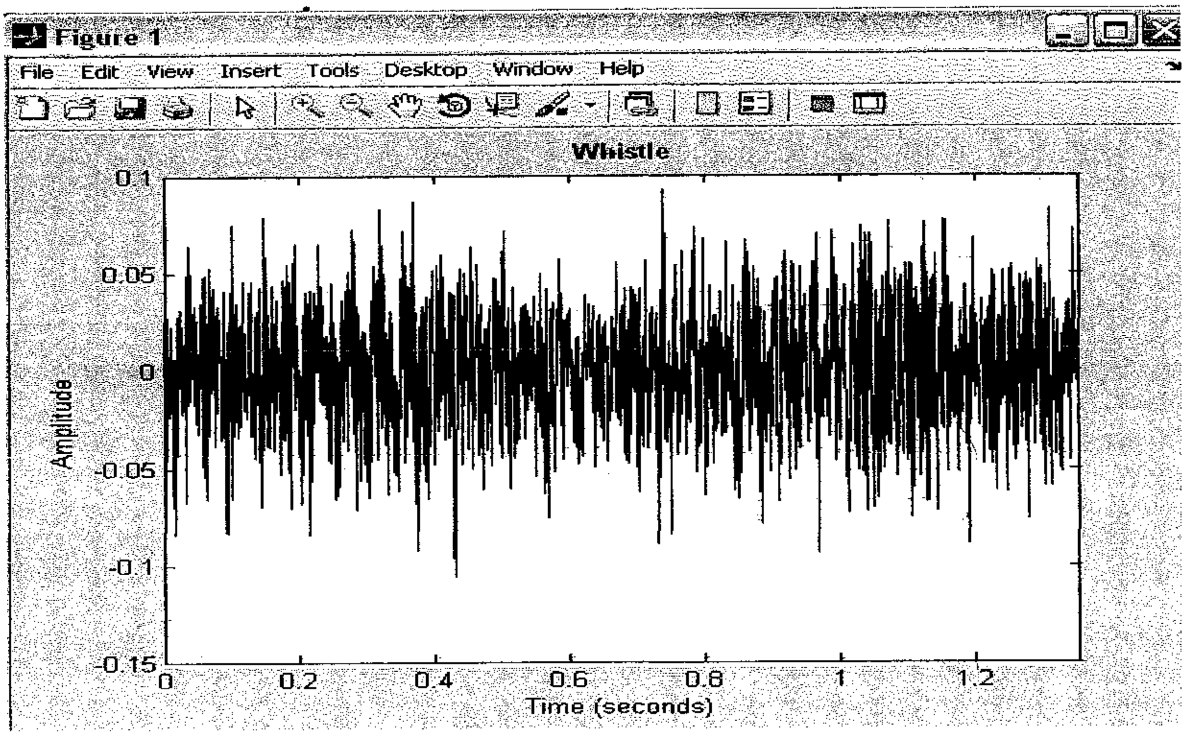


Figure 11: The amplitude over time of the whistle audio file

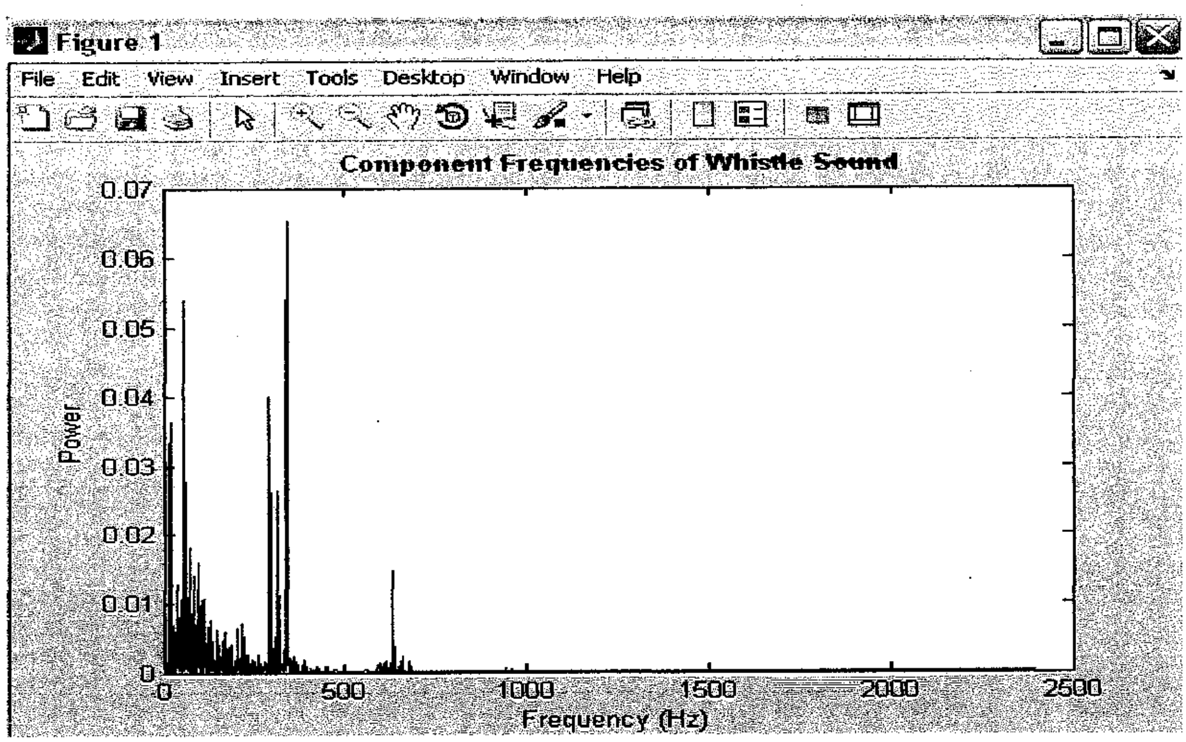


Figure 12: The FFT of the whistle sound

The code below has been used for the first analysis by using MATLAB, further improvement and development will be performed in order to produce a better codes and spectral analysis of the audio files.

```
[x,fs] = wavread('FYP_Audio_0001.wav');

plot(x)
xlabel('Sample Number')
ylabel('Amplitude')
title('{\bf Football Playback}')

sound(x,fs)
firstportion= x(2.45e4:3.10e4);
t = 10*(0:1/fs:(length(firstportion)-1)/fs); % Time base

plot(t,firstportion)
xlim([0 tb(end)])
xlabel('Time (seconds)')
ylabel('Amplitude')
title('{\bf Football Playback}')

m = length(firstportion); % Window length
n = pow2(nextpow2(m));  % Transform length
y = fft(firstportion);  % DFT of signal
f = (0:n-1)*(fs/n)/10;  % Frequency range
p = y.*conj(y)/n;       % Power of the DFT
%Plot the first half of the periodogram, up to the Nyquist
frequency:

plot(f(1:floor(n/2)),p(1:floor(n/2)))
xlabel('Frequency (Hz)')
ylabel('Power')
title('{\bf Component Frequencies of Football Playback}')
```

Spectrum Analysis 2: FFT by using Audacity software

MATLAB is a good tool in analyzing the audio file. However, there are more accurate and widely known software for FFT which are the Audacity and Sound Forge. The Author performed the Spectral Analysis by using the same three audio files in both Audacity and Sound Forge software so that the actual, accurate and average values or results can be obtained to further use it for matching algorithm. The analysis can be seen as below:

Figure 13: The spectral analysis of the full playback

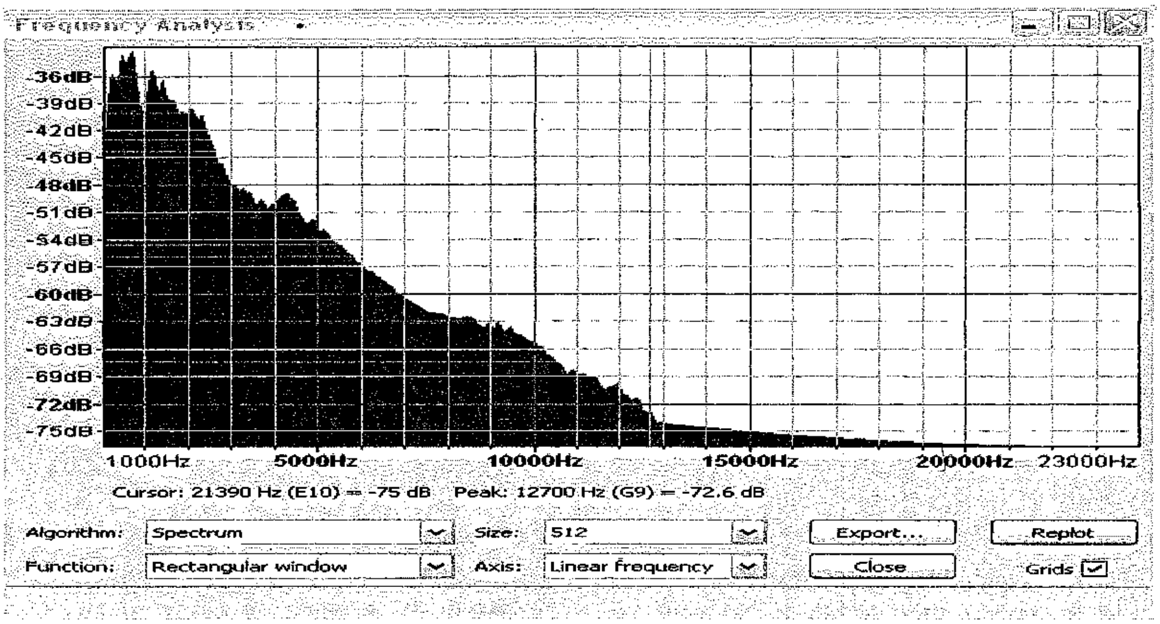


Figure 14: The frequency analysis of the full playback

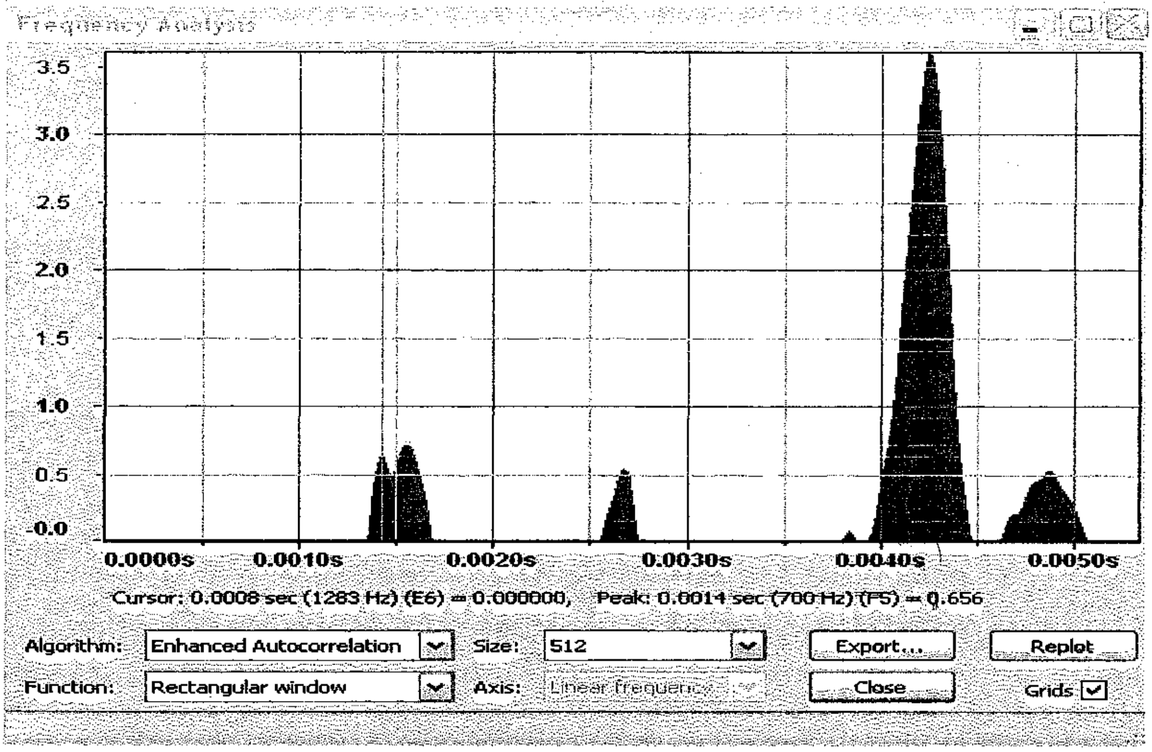


Figure 15: Autocorrelation functions to see it clearer

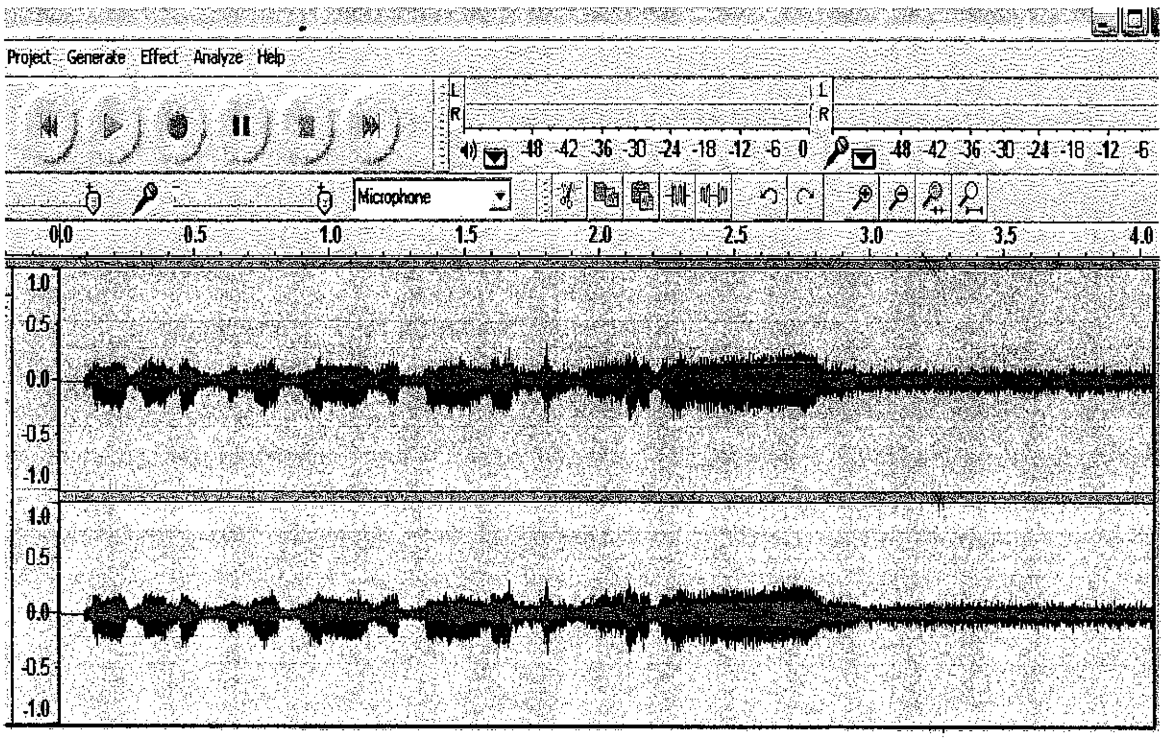


Figure 16: The amplitude of the goal sound

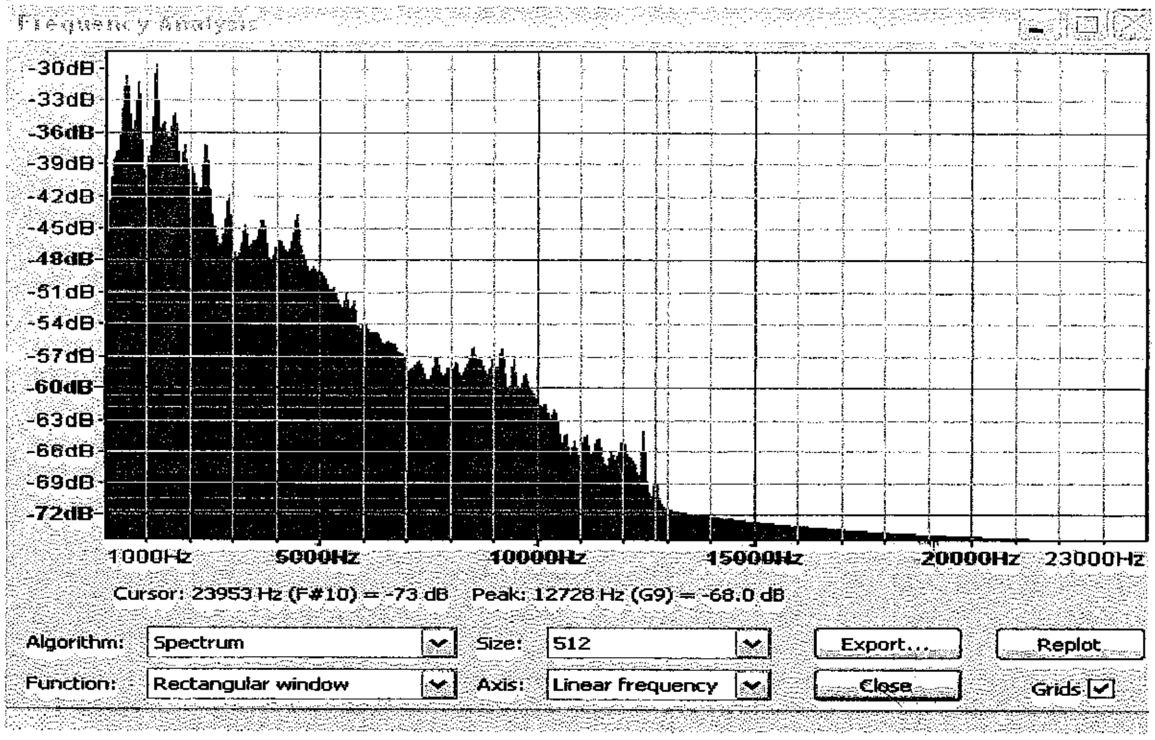


Figure 17: The frequency analysis of the goal sound

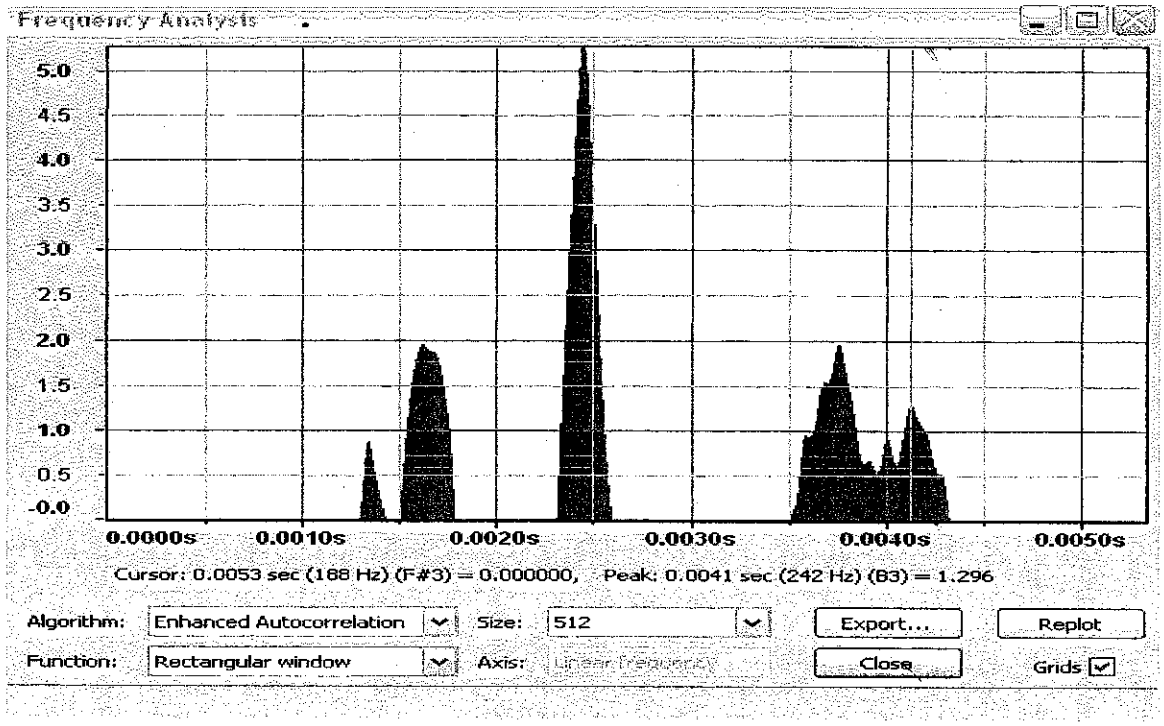


Figure 18: The enhanced autocorrelation view

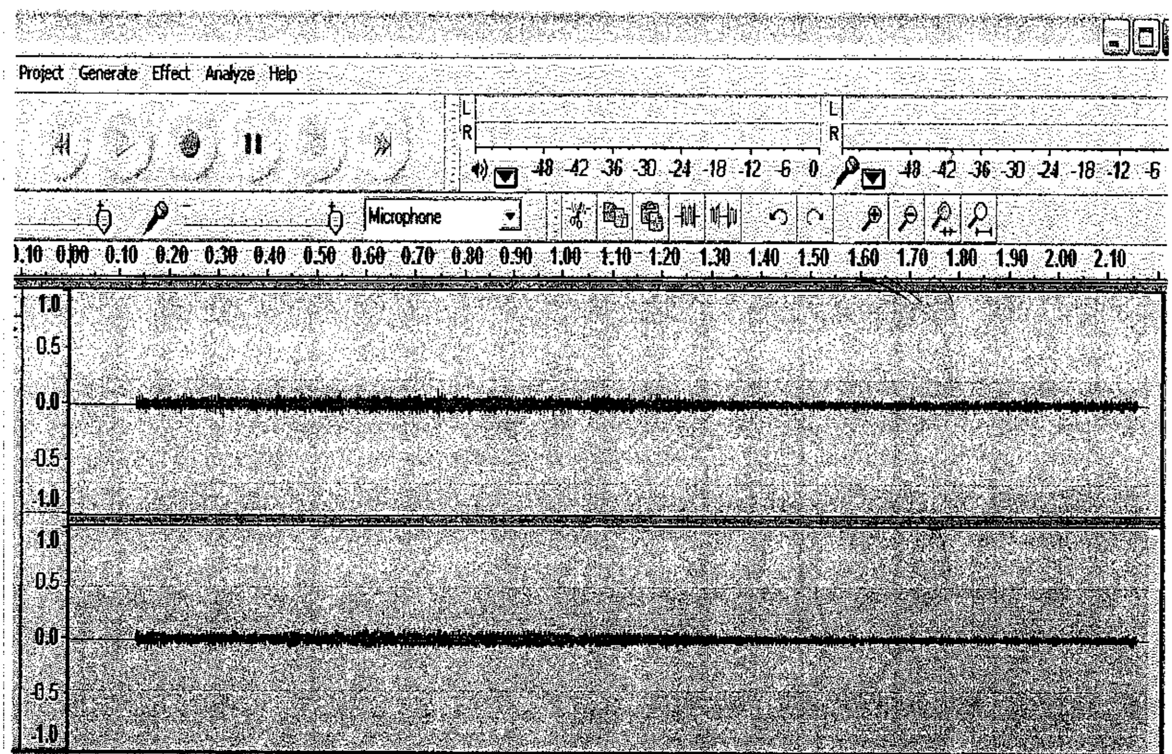


Figure 19: The analysis of the whistle sound

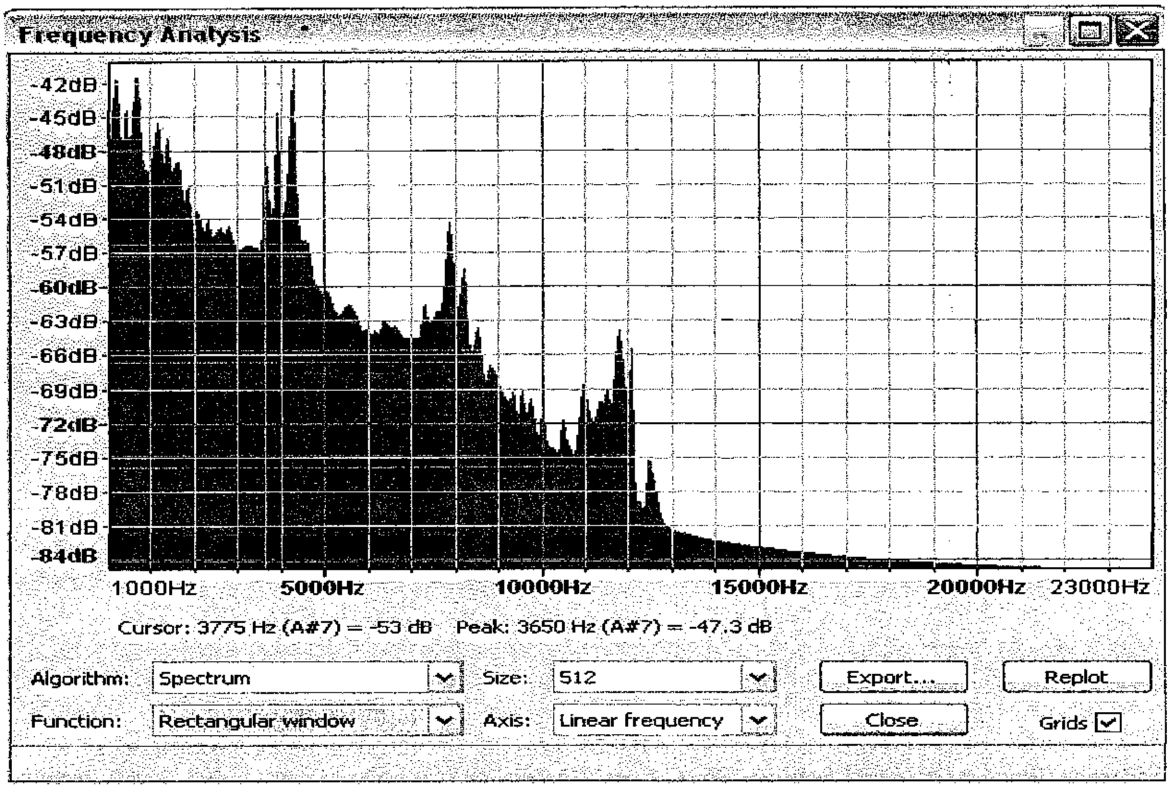


Figure 20: The frequency analysis of the whistle sound

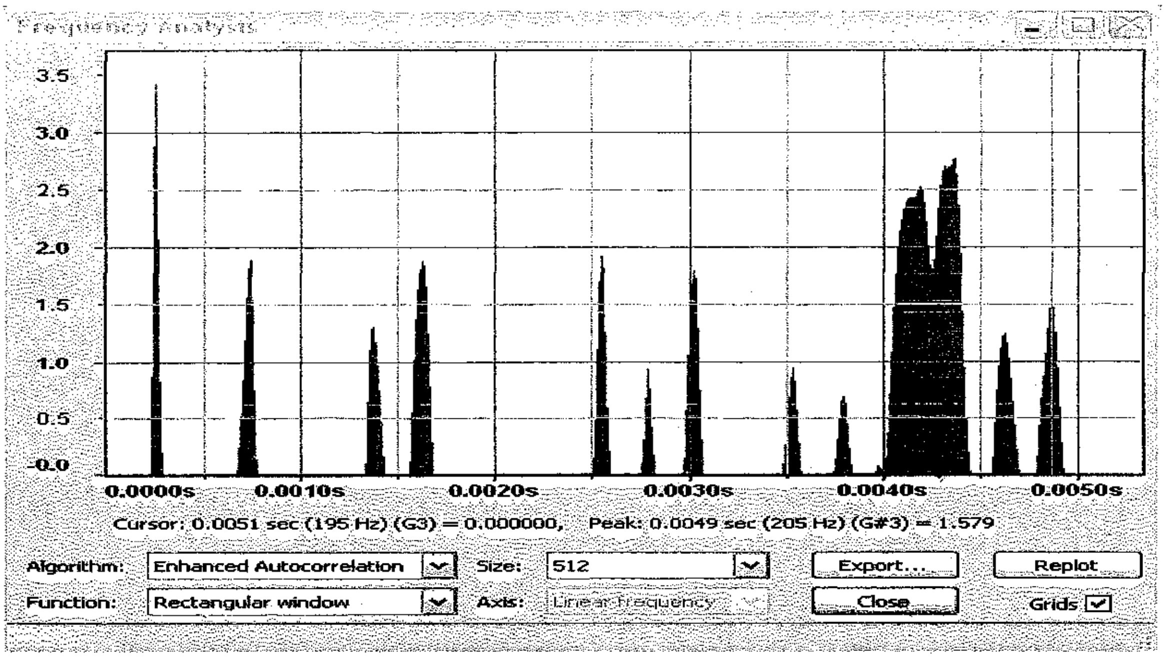


Figure 21: The autocorrelation view of the whistle sound

Spectrum Analysis 3:-FFT by using Sound Forge software

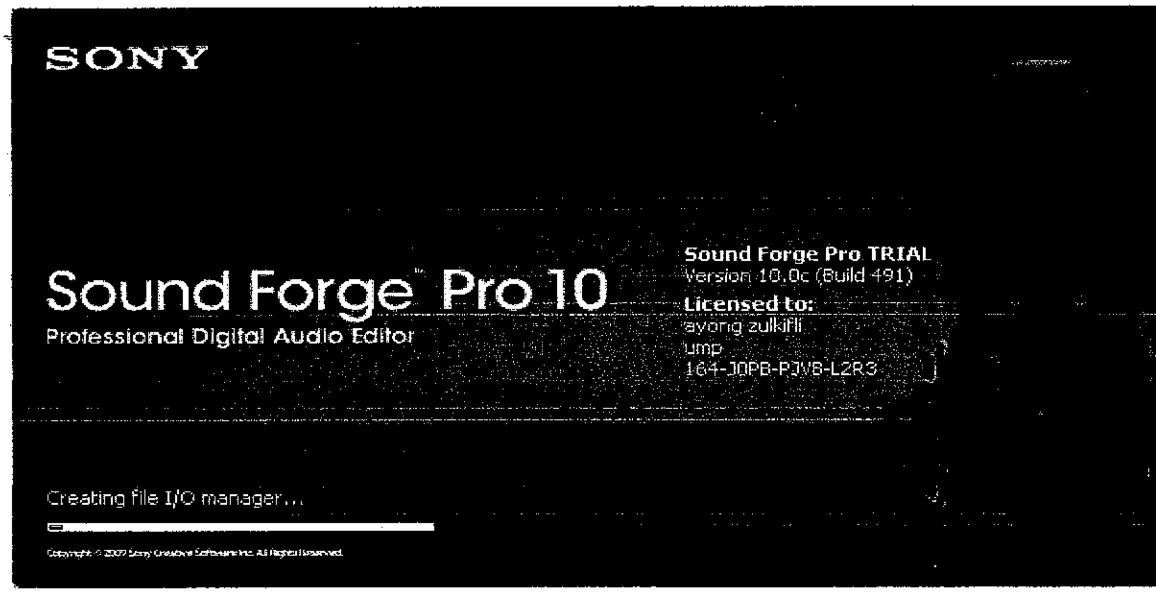


Figure 22: Sound Forge Software from SONY

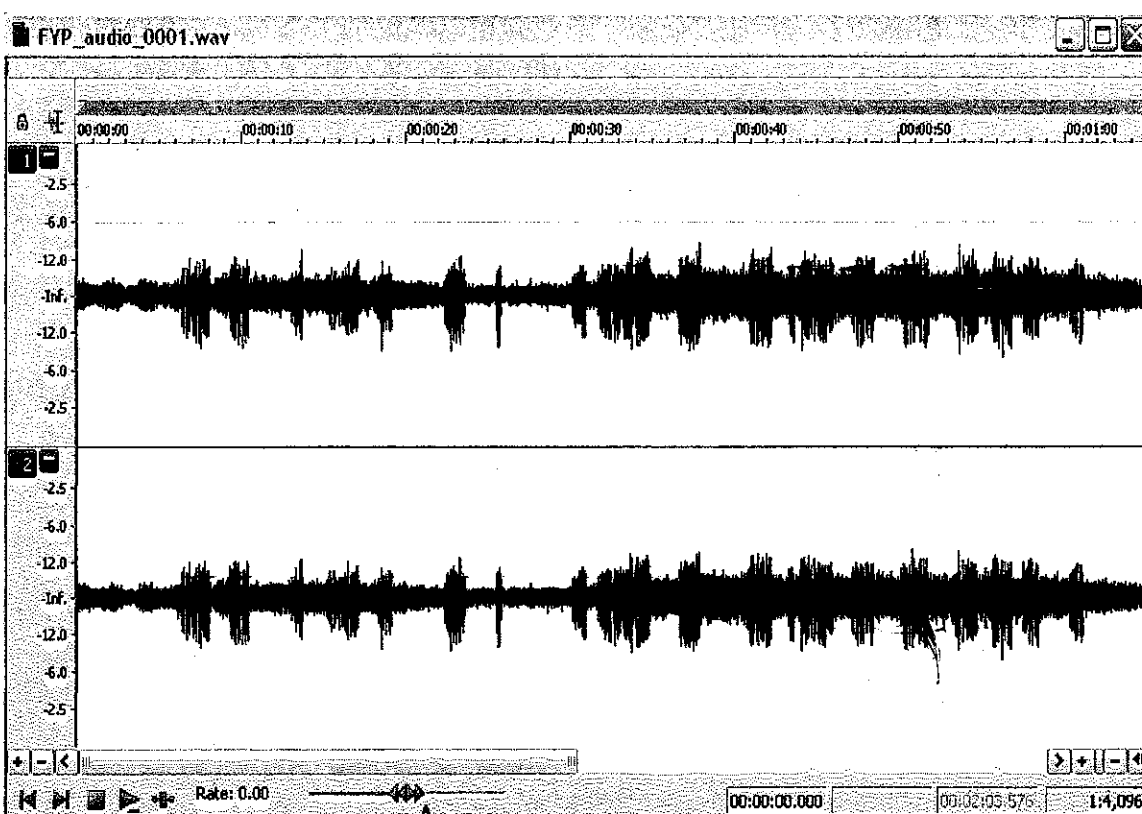


Figure 23: The spectral analysis of the football playback

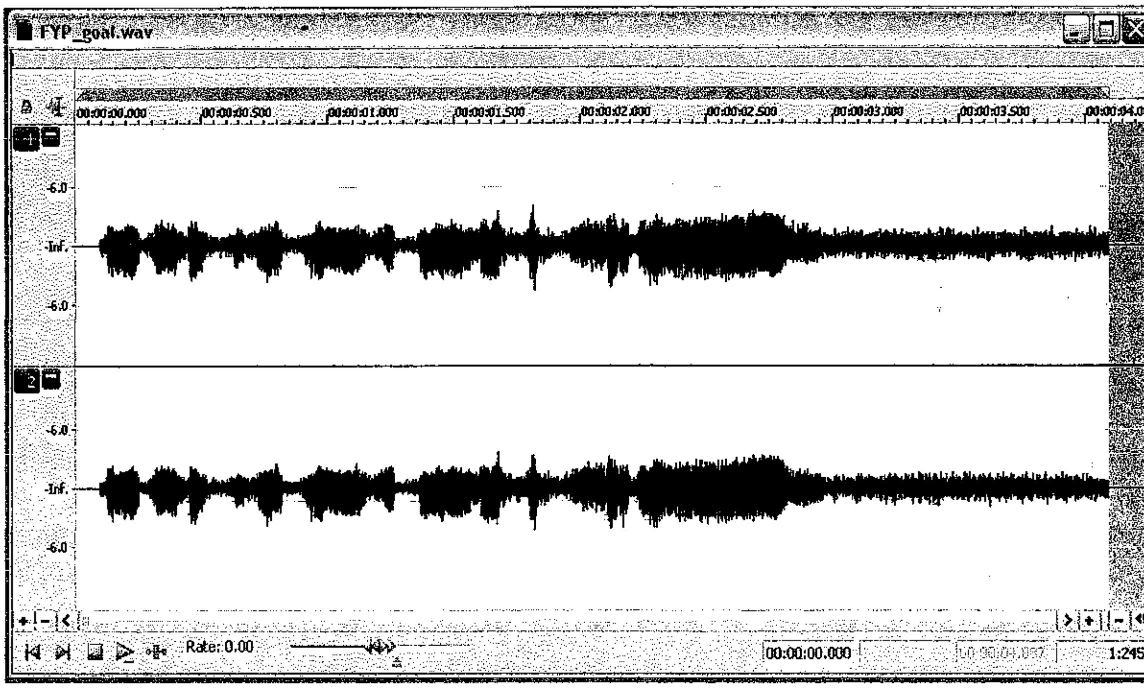


Figure 24: The analysis of the goal audio file

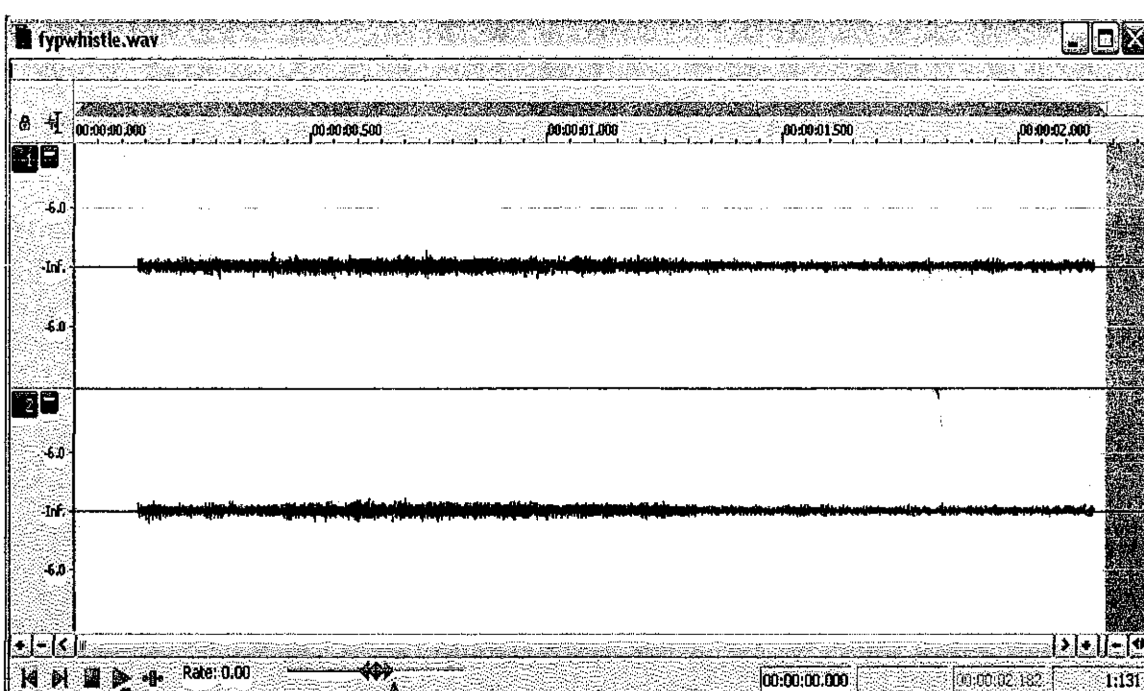


Figure 25: The analysis of the whistle audio file

One of the features of the Sound Forge Software is where the users can perform the spectrogram analysis of the audio file. As shown below, all the three audio files were being analyzed and plotted in one graph. This is where we can see the intersections of the frequency of goal and whistle sound with the full playback. It differs by the color of the line on the graphs. There are certain value of the property of sounds can be obtained from the process. Further detailed analysis of the values will be done from time to time.

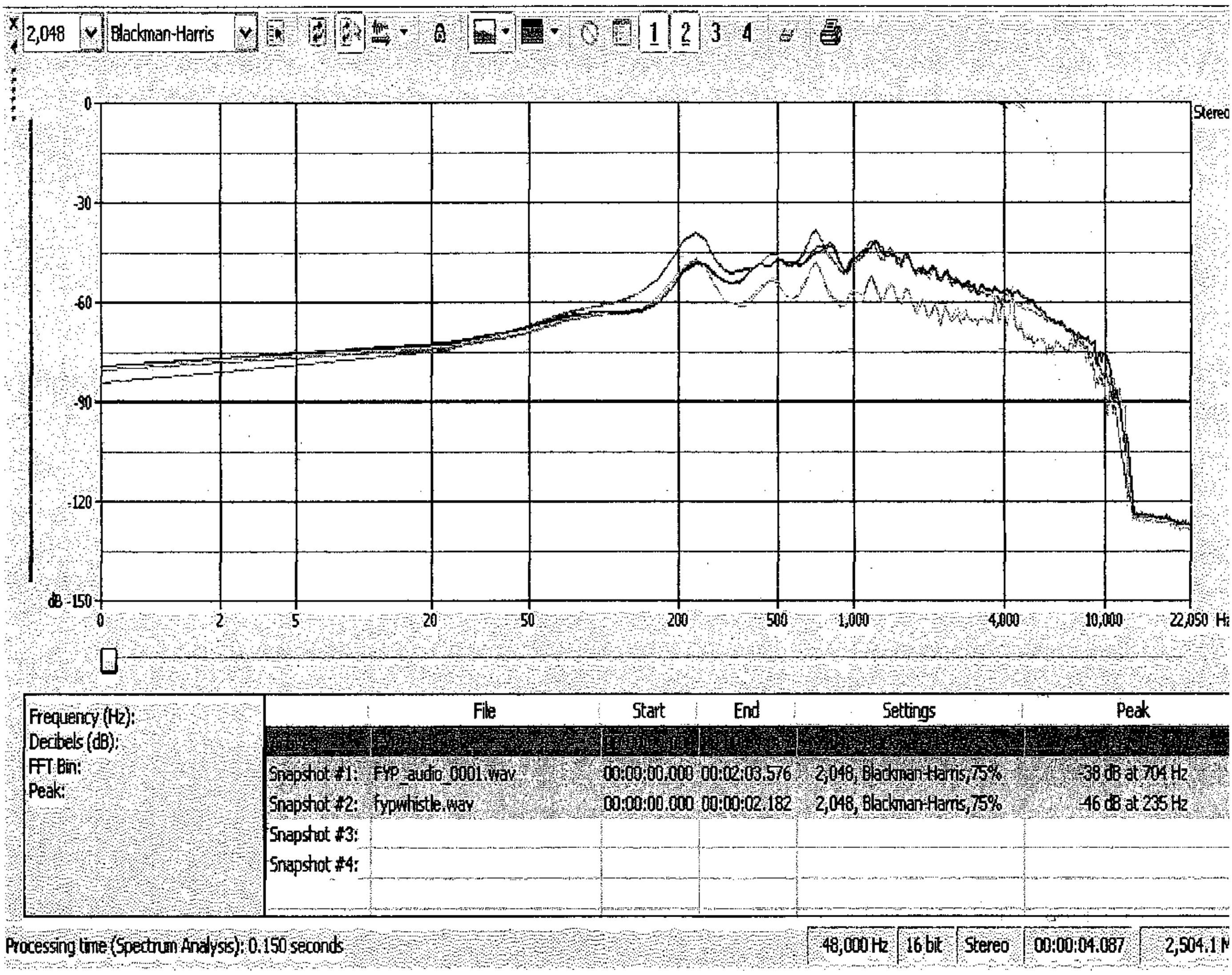


Figure 26 : The combination of three audio files

As there are lots of noises in the football playback, the author analyze and reduce the noise by using both Audacity's noise removal and MATLAB coding for noise removal. The difference of the sound can be seen as below:

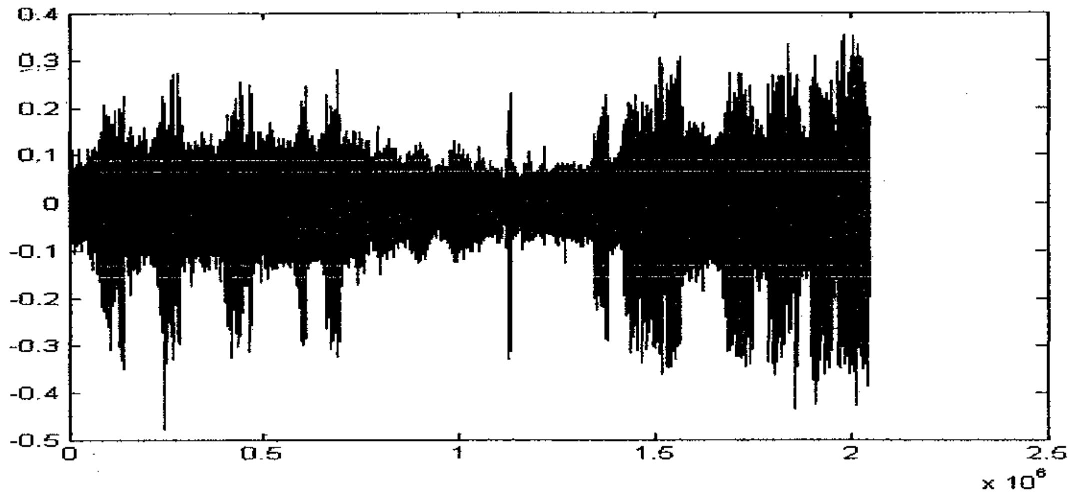


Figure 27 : Playback: Before noise filtering

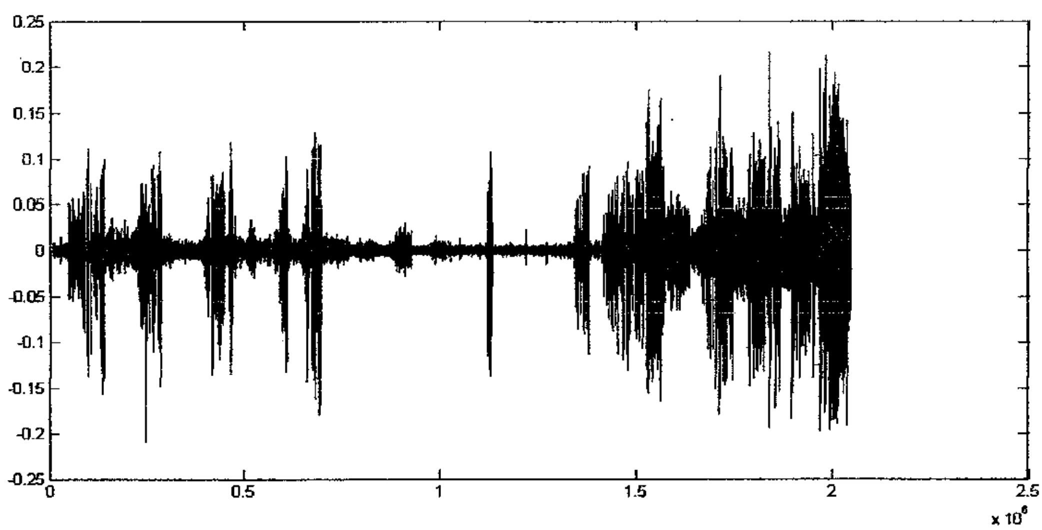


Figure 28: Playback: After noise filtering

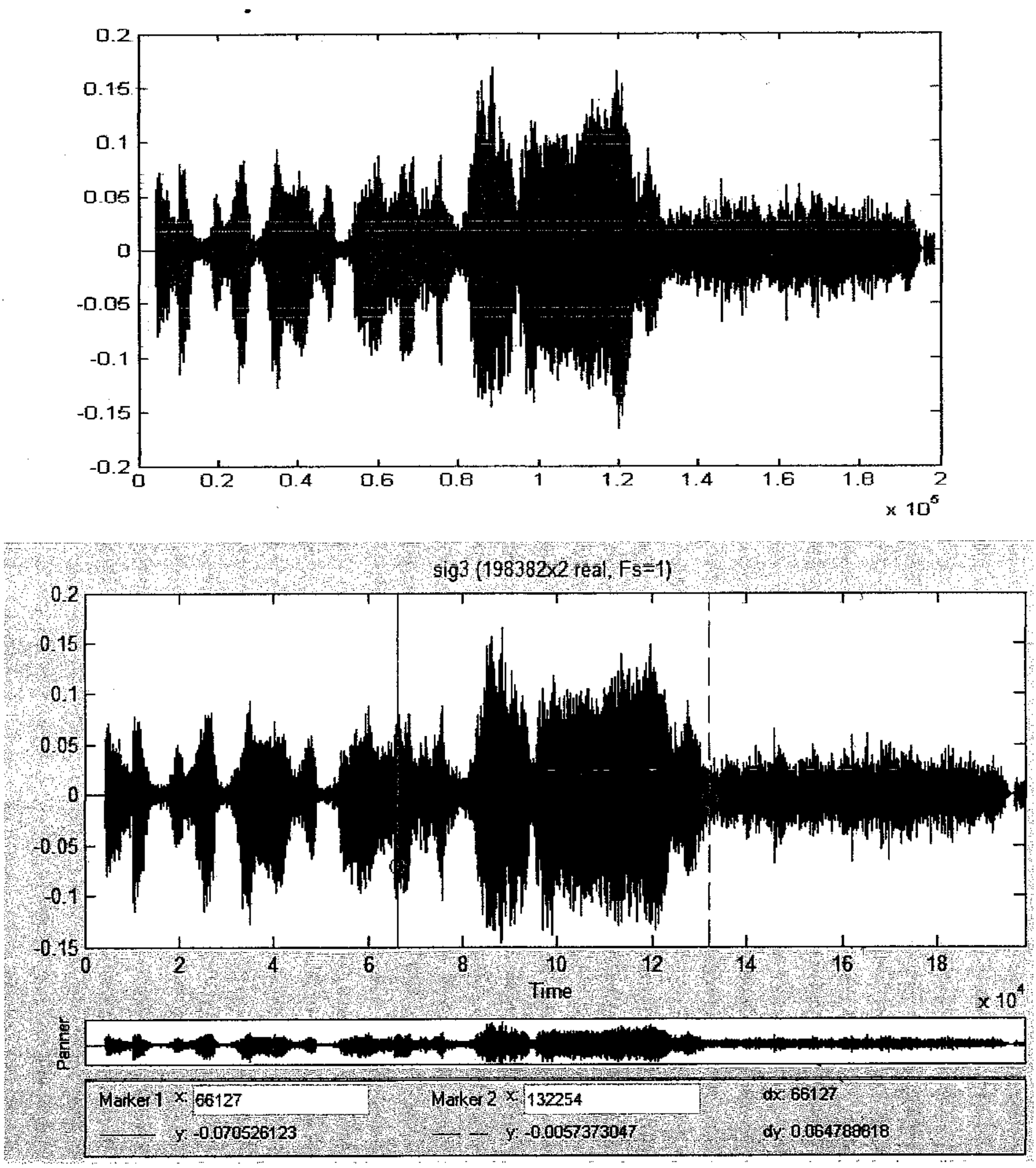
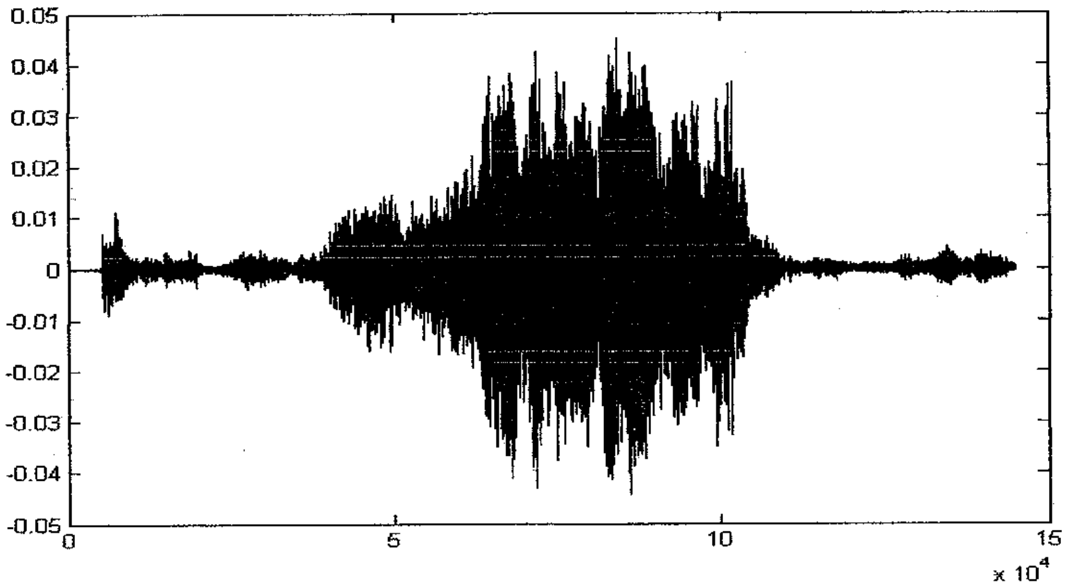


Figure 29 & 30: The goal sound after noise



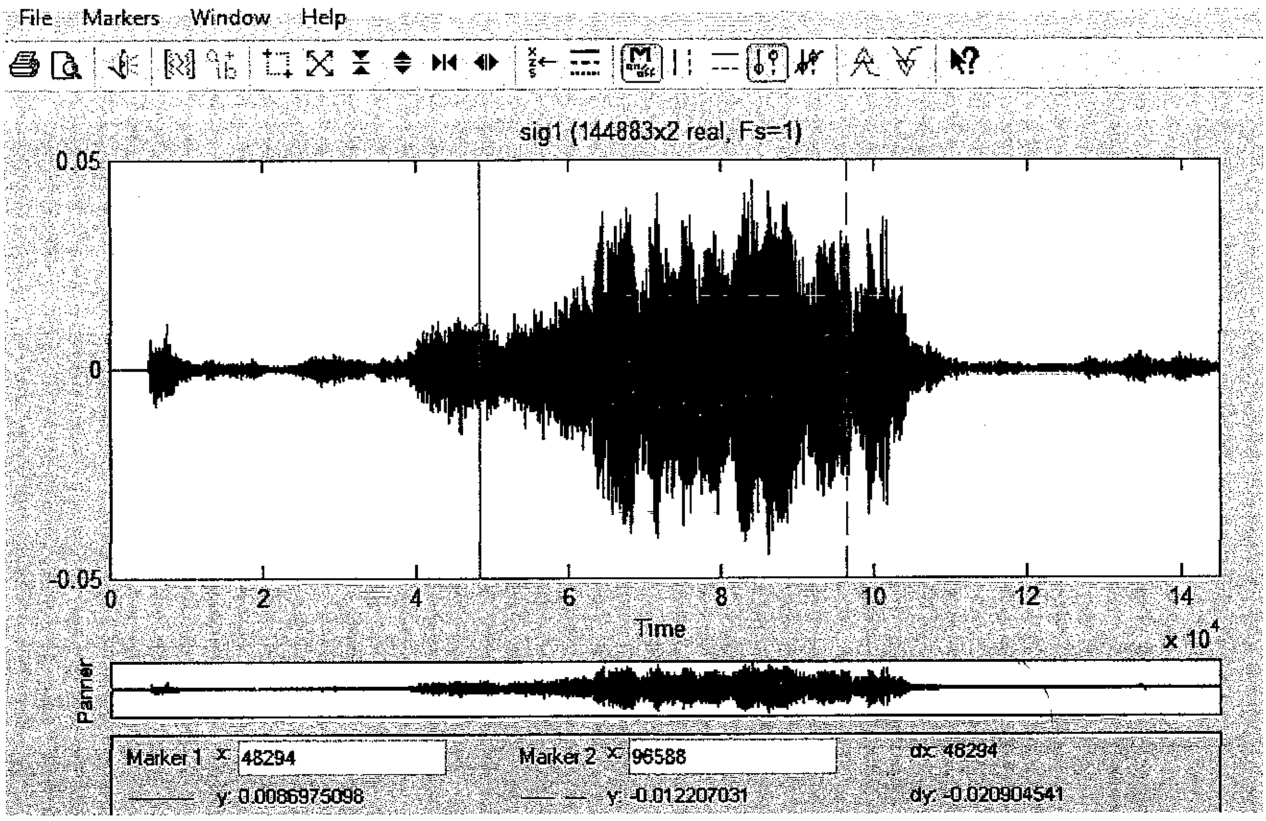


Figure 31& 32: Whistle sound by using SP Tools and after filtering the noise

In general, there are 4 types of filtering for noises which are [42]:

1. **Low pass filter**

   This filter is suitable to be used when the sound is not really audible to hear or in other words, the low frequency signals. It allows or passes the low frequency signals but attenuate or reduce the amplitude of signals with frequency higher that the cut off frequency, which being defined by the analyzer .

2. **High pass filter**

   This filter is the opposite of low pass filter where it passes the high frequency signals but attenuate the low frequency signals.

3. **Band pass filter**

   It is the combination for both low pass filter and high pass filter .

4. **Stop band filter**

   The opposite of band pass filter where it passes most frequencies unaltered, but attenuates those in a specific range to very low levels .

For the matching part, the author is using band pass filter of least squares linear-phase band pass finite response (FIR) filter. The author focused on the whistle sound rather than goal sound as it is easier to analyze as the goal sound is differ for each goal because the commentator usually shout the name of the player that scored, not the word 'goal'. The whistle sound can be considered as a high pitched sound which is different from other sound such as commentator speech, audience applause and so on. The cut-off frequencies of the band pass filter are chosen based on the spectral characteristics of the whistle sound which has been perform earlier. From the spectral analysis, the author gets to know the lower and upper frequencies of the spectral components of the whistle sound. The author also compared the original sound of whistle as well as 7 other whistle sounds from various football playbacks to get to know the mean of frequency range of whistle. In average, the frequency of the whistle sound is between 2000 to 2400 Hz. The analysis can be seen as below:

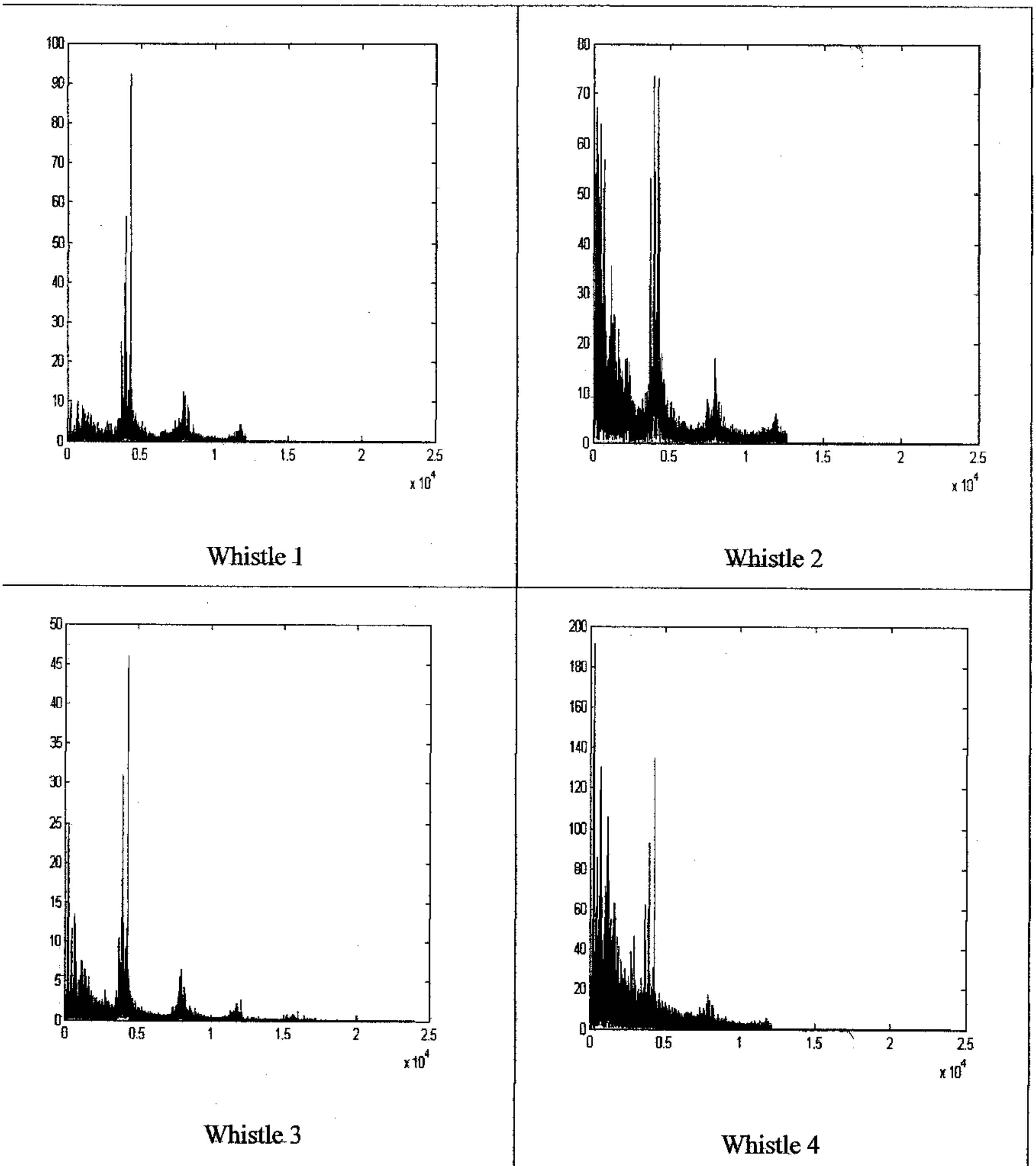


Figure 33: Whistle audio sample

4.2 Discussions

- **Brain Processing**

According to Health Scan, the blind people use the same brain region as sighted people in reading [38]. The part of the brain responsible for visual reading doesn't require vision at all, according to a new study by researchers from the Hebrew University of Jerusalem and France. Brain imaging studies of blind people as they read words in Braille show activity in exactly the same brain region that lights up when sighted people read. However, as the team has interviewed Mr Moses from the Malaysian Association for the Blind (MAB), he claimed that the blinds are usually very passive to mix and learn new things unless if they have the interest on that area. As for he himself, he managed to read, use the keyboard and recognize well on the environment around him because of practices and learning in daily life [39].Hence, as the characteristics of every person is different, it can be conclude that the brain processing of the blind people is the same as sighted people, depends on how well the blind people managed to adapt and learn new things. However, for this project, by focusing on the second group of blind people, more or less, they already have the advantage as they had seen the football match before, it can be used by the blind people since birth, but will definitely take more time to adapt than the second group.

- **Noise**

Any project which involves sounds and audio analysis, noise is one of the typical problems that usually happen. However, the noise for football playback or discrete sampling is not as greater as for the live match. The measurement of noise levels depends on the bandwidth of the measurement. Discrete frequency components theoretically have zero bandwidth and therefore it should not be scaled with the number of points or frequency range of the FFT. To compute the SNR, compare the peak power in the frequencies of interest to the broadband noise level. This is really crucial as to determine the frequencies of certain sound, for example the sound of whistle and goal, it will need to display the accurate value of frequencies so that it can be matched with the stored sound frequencies thus produce the vibrations. With the noise, this will cause distortion and inaccurate frequency values.

- **Aliasing**

According to Nyquist theorem, a signal must be sampled at a rate equal to twice that of the highest frequency. Aliasing happen when any frequency components higher than the Nyquist rate which will appear in the measurement as a lower frequency component. To avoid aliasing, a low pass filter is placed ahead of the sampler to remove any unwanted high frequency elements. Figure 3 shows an adequately sampled signal and an under sampled signal. In the under sampled case, the result is an aliased signal that appears to be at a lower frequency than the actual signal.

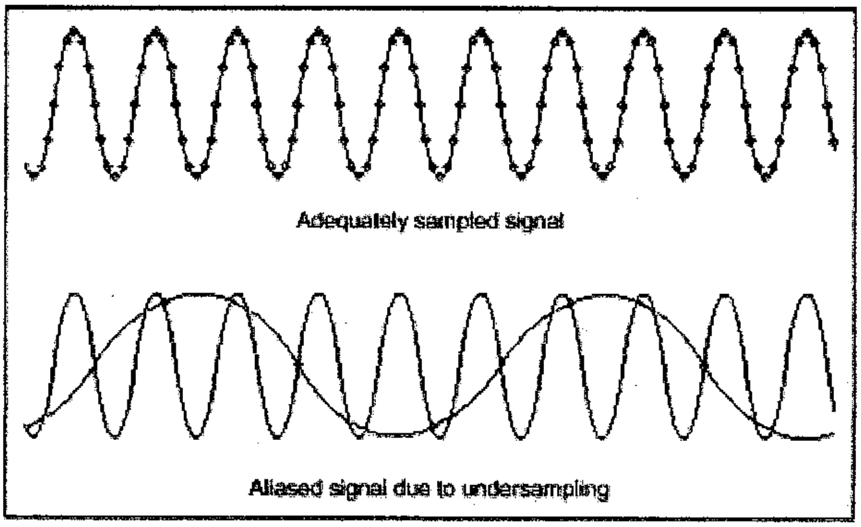


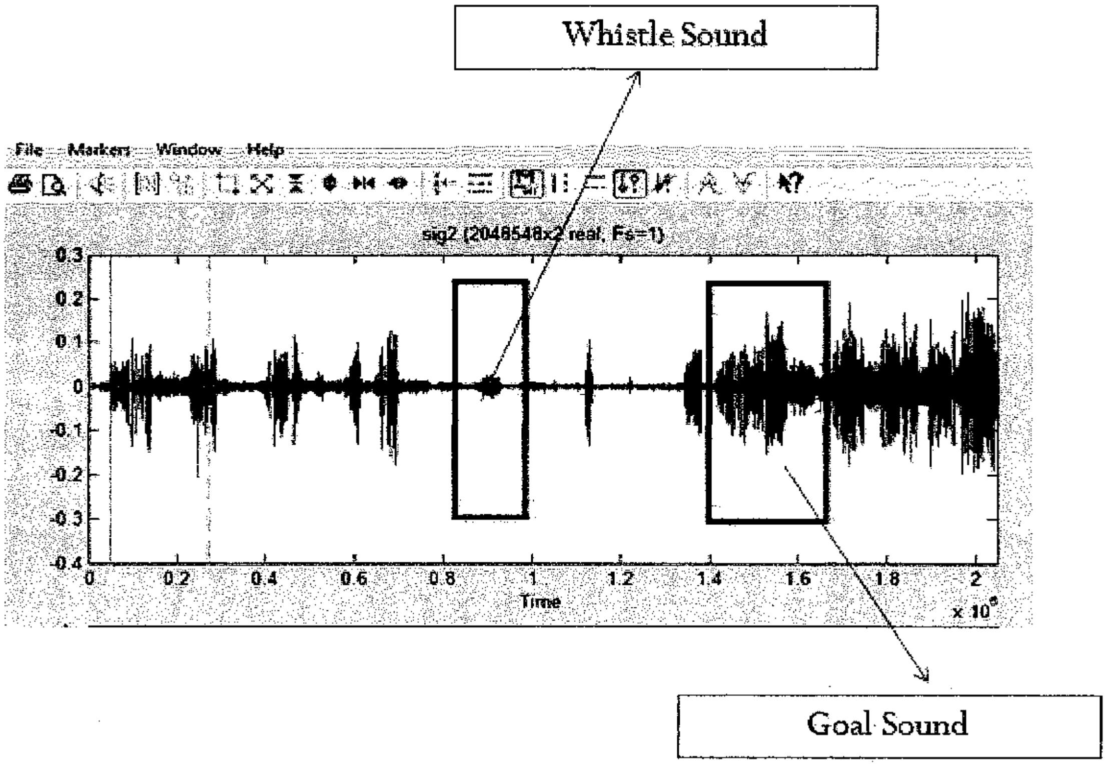
Figure 34: Adequate and Inadequate Signal Sampling

- **Sound Matching**

For the sound matching part, the author faced challenges in finding the exact sound pattern or value of the whistle as the sound spectrum of the football playback is very wide with lots of noise in it. However, the step for the matching part can be concluded as below:

- The sound of playback is being sampled and segmented into frequencies, amplitude, and power.
- The FFT of the sound is being calculated to get the exact or the nearest accurate sound value for both playback and whistle.
- As there are lots of background noises from the playback, it is very hard to identify or detect the value of whistle sound in the full playback. Thus, the author performs noise filtering to remove all the noise.
- The author still needs to get a cleaner or to filter the sounds in the playback by using threshold. The value of the sampled sound should be bigger than the threshold which indicates that the value of whistle is high pitch. The samples are being sort out from the large playback sound from each step.
- Next, the author needs to perform the Short-Time Energy to estimate in order to detect the whistle sound segment. The value is then being compared.
- The steps iterate until the value of whistle is matched with its corresponding segments in the full playback.

The author is having difficulties in matching the sound. It was a very challenging task which further time and advanced processor and tool need to be explored. After consulting with the expertise, the steps are correct but and yet, we are still working on it, to resolve the problems and issues. The issues are discusses further at the next page.



Most of the steps listed have been done by the author. Each segments in the playback is being extracted into N window length to calculate the amplitude and frequency of it, so that the matching part can be done when the specific or range value of the sound is known. However, it is very challenging when it comes to deal with the signal processing. You will need a powerful processor as well as accurate matching algorithm as the spectrum is wide, thus it is very hard to detect where the whistle segment is. The figure below shows the expected output of the audio files after the matching algorithm take place. The green color is referring to the whistle sound while the red color is the place where the goal sound resides. When it matched the whistle sound in green, then these indicates that, this is the place where the vibrations will take place. Same goes for the goal sound. Typically, vibrations will occur when the system detect the frequency and by using timer. This is a major drawback as usually it will vibrate not according to meaningful events like goal and whistle. Thus, the author is suggesting vibrations at meaningful events, with the help of audio processing to detect where the meaningful sound events are. Mind that, the scope of this project is only to analyze the sound and perform the matching part. The author is working with other two members, thus the vibration task will be continued by the third person, Then when we combined, we will able to make the Immersive Assisting tools and Application for the Visually Impaired People.

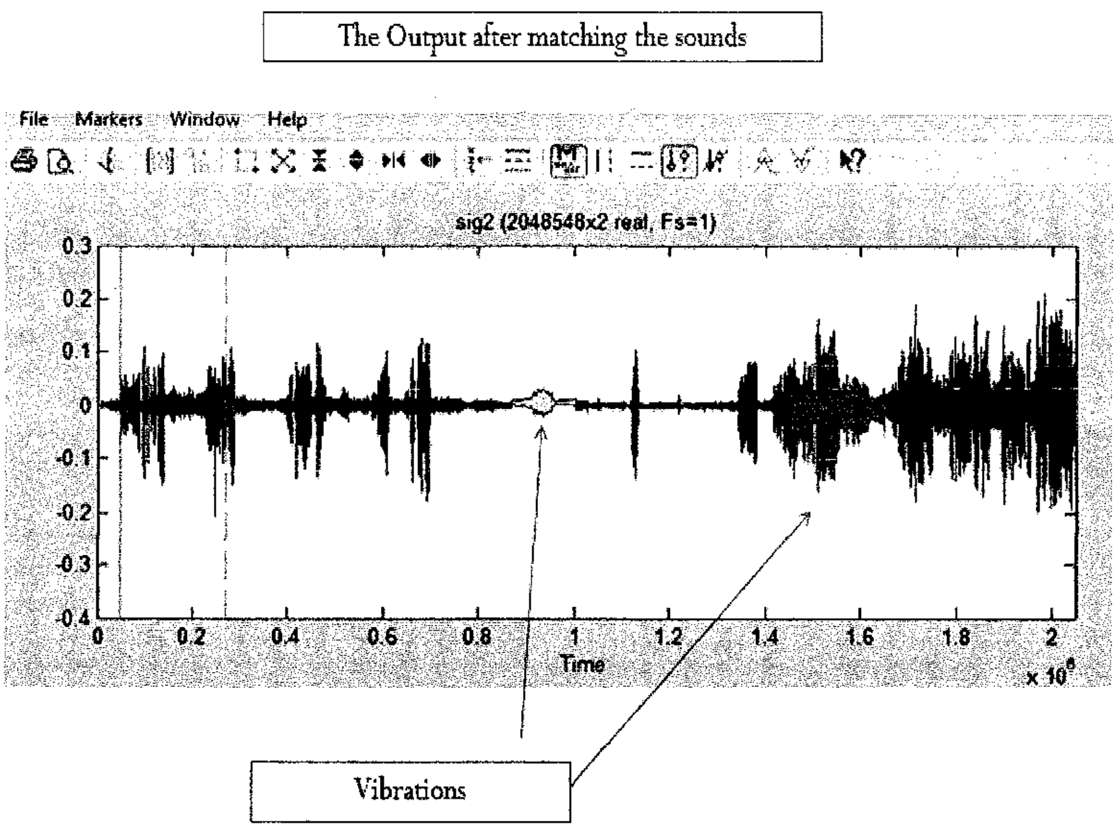
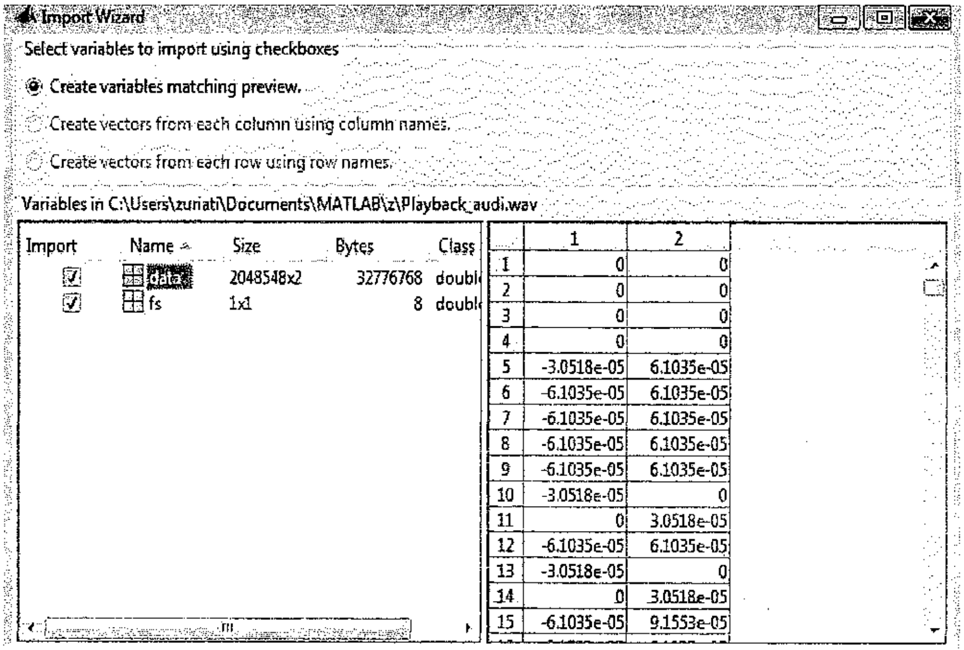
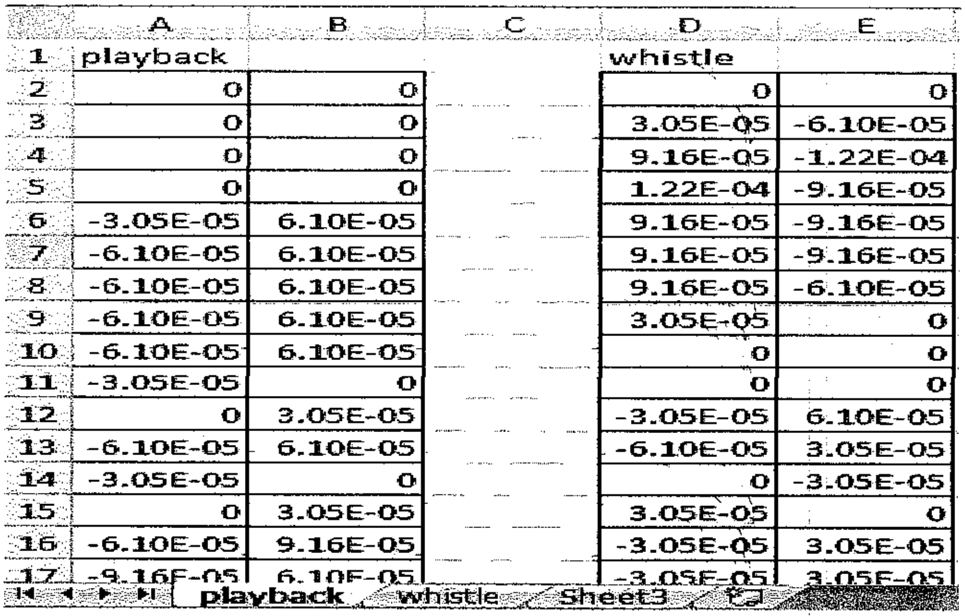


Figure 36 and 37 below is the data of the audio sound for both playback and whistle. The author is trying to show the challenges in getting the accurate data. As can be seen below, the data values are very much the same , thus it is really hard to compare and find the intersects of whistle sound data value with the football playback data value.



Import	Name	Size	Bytes	Class
☑	data	2048548x2	32776768	doubl
☑	fs	1x1	8	doubl

	1	2
1	0	0
2	0	0
3	0	0
4	0	0
5	-3.0518e-05	6.1035e-05
6	-6.1035e-05	6.1035e-05
7	-6.1035e-05	6.1035e-05
8	-6.1035e-05	6.1035e-05
9	-6.1035e-05	6.1035e-05
10	-3.0518e-05	0
11	0	3.0518e-05
12	-6.1035e-05	6.1035e-05
13	-3.0518e-05	0
14	0	3.0518e-05
15	-6.1035e-05	9.1553e-05



	A	B	C	D	E
1	playback			whistle	
2	0	0		0	0
3	0	0		3.05E-05	-6.10E-05
4	0	0		9.16E-05	-1.22E-04
5	0	0		1.22E-04	-9.16E-05
6	-3.05E-05	6.10E-05		9.16E-05	-9.16E-05
7	-6.10E-05	6.10E-05		9.16E-05	-9.16E-05
8	-6.10E-05	6.10E-05		9.16E-05	-6.10E-05
9	-6.10E-05	6.10E-05		3.05E-05	0
10	-6.10E-05	6.10E-05		0	0
11	-3.05E-05	0		0	0
12	0	3.05E-05		-3.05E-05	6.10E-05
13	-6.10E-05	6.10E-05		-6.10E-05	3.05E-05
14	-3.05E-05	0		0	-3.05E-05
15	0	3.05E-05		3.05E-05	0
16	-6.10E-05	9.16E-05		-3.05E-05	3.05E-05
17	-9.16E-05	6.10E-05		-3.05E-05	3.05E-05

Figure 35&36 : Playback and Signal data

The figures below shows the Fast Fourier Transform of the football playback and the whistle sound. Firstly, the power of the sound is slightly different as full playback has a higher power as compared to whistle sound. Meaning to say, as the power of the whistle sound is lower, these indicate that the whistle sound is very small and tiny inside the FFT of the foot ball playback. Thus, these has becomes a big challenge to the author in order to do the matching part. Further effort and studies will be conducted in order to fulfill all the objectives listed for this project provided with adequate time and resources.

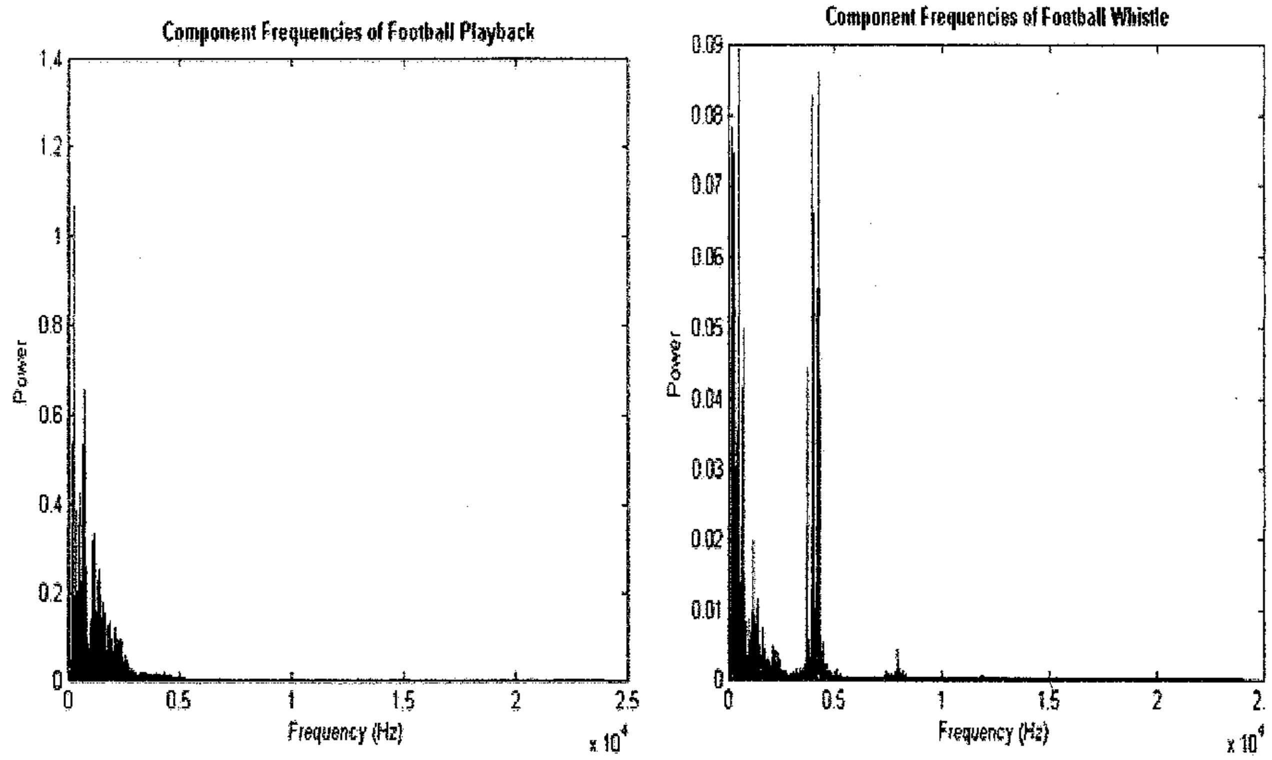


Figure 37: FFT comparison between playback and whistle

CHAPTER 5

CONCLUSIONS AND RECOMMENDATIONS

This project highlights the design and development of the Immersive Assisting Tool and Application for the visually impaired people. By focusing on Audio, which to analyze the frequency of the audio file to generate different types of vibrations in order for the blind people to feel and 'watch' the football game. Audio is proven to be a really helpful sense for the blind as it gives the auditory stimuli and queue for navigation, as well as perception on something. With the advancement of technology, the objective of this project can be done successfully with more findings to improve the system. At the end of day, all of these three main components which are the Audio, Video and Haptics will be combined, thus producing a complete set of Immersive Assisting Tools and Application for the Visually Impaired People.

The author managed to fulfill all the objectives of the research, except for the matching part. There are many works and scientific calculations involved in the area of signal processing which usually under the Electric and Electronic Engineering department. The author need to study, research and put more effort in order to complete the matching part as it is a complex process, even more challenging than image recognition or processing. The author has learned a lot of lesson throughout this project:

- **To be independent**

  Final year project is about our own project, it is not the same as group project or course project that typically every student has gone through. The author learned to be independent and not to depend much on the supervisor. It should be that way, as the task of supervisor is to supervise not to completely help the students and assist them one by one. Hence, it is a good experience as this project is your own responsibility, you are the one who will do it, conduct research, present and defend it.

- **Increases Level of Communication Skills and Confidence**

  The author has gone through lots of presentations since in the first year. However, through this project, the author has been exposed to exhibition presentation style through poster/ pre-SEDEX presentation. SEDEX stands for Science, Engineering Exhibition where most of researchers presents their findings. Hence, during pre-SEDEX, the author was selected to go to SEDEX and present in front of the external examiner. Besides that, the author also managed to enhance the communication skills through VIVA or proposal defense.

- **Time Management**

  The author realized the need to arrange the task properly and plan for the project activities through Gantt chart and follow it consistently. In a project, the first phase is to plan about what to do, when to do, where to do, who is the responsible person to do and how to do. Good time management is one of the key factors for a project to be successful.

- **Leadership**

  This is one of the key components in doing a project. In fact, everyone is a leader to oneself.

This area of studies has a lot more potential to be explored. For future work expansion and continuation, the author suggests and recommends the project to be further improved on:

- **Research for live football matches**

  From the continuation of FIFA WorldCup 2010, where the blind was being equipped with headphones, seats and personal commentator and watching live football matches, this project should be able to process the live signal. It would be fun and exciting if the blinds can watch it live by using the tools and immersive application from this project as this project process the discrete signal from football playback. Hence, various issues such as synchronization between the live signal and the system, live noise, how to detect the game match, accessibility, availability, power processing need to be further explored.

- **Translate the image captured by camera into frequencies and sound**
  Besides that, the project also can be further improved by using the high resolution camera technology in capturing the data on the football matches thus translating the current situation to sets of frequencies which then can be use to tell the blinds on what actually happened or even process it to produce vibrations.

- **Use sensors to respond with changing and the excitement of the live game**
  As the blind people are being equipped with the seats/ special chair, sensors can be use in helping them to immerse with the game. For example, when one of the players kick the ball by using left foot, then the sensors will respond and vibrate on the left foot of the blind (on the seat). This is one of the way, the blind can experience the game and get the excitement and intense from the sensor's chair.

- **Use speech synthesizer for user navigation**
  Speech Synthesizer is really a powerful tool for the blind people. It takes the input from the blinds voice thus gives appropriate response in navigation part of the systems. Speech Synthesizer can be done by using MATLAB 7.

- **Game analysis and coordination with the movement of players and ball**
  To make the live signal from the football match synchronize with the live signal to the device, deep research need to be done. The game analysis need to be done in order to know the coordinate of the ball or players, so that their position can be tracked and showed in the device. Besides that, the researcher should also consider the movement of the ball/player, if they are out from the football match. Use for other activities/sports, for example, tennis, badminton as well as watching movie (cinema), learning from playback , playing games and more.

REFERENCES

[1] Saito, H., Inamoto, N., Iwase, S," Sports scene analysis and visualization from multiple view video", IEEE International Conference on Multimedia and Expo (ICME), vol.2.pp.1395-1398,2004

[2] Y.You, H. Lee, MY.Sung, K.Jun, JS Kang, "Sound-Specific Vibration, Its Performance of Tactile Effects and Applications"The 9th International Conference for Young Computer Scientis, University of Inchieon, Republic of Korea

[3] http://www.who.int/mediacentre/factsheets/fs282/en/

[4] http://www.brailleinstitute.org/facts_about_sight_loss

[5] http://www.mab.org.my/

[6] http://www.mab.org.my/reference/blindness.html

[7] http://www.mydigitallife.co.za/index.php?option=com_content&vie=article&id=1054651:everyones-world-cup&catid=18:home-a-entertainment&Itemid=251

[8] A Multimodal Interaction Scheme Between a Blind User and the Tyflos Assistive Prototype
Nikolaos Bourbakis1,3, Robert Keefer1, Dimitrios Dakopoulos1 and Anna Esposito1,2 *1Wright State University, ATRC, Dayton, OH 454352 Second University of Naples, Naples, Italy*

[9] Mobile audio navigation interfaces for the blind

[10] A Mobile Audio-Visual Terminal for the DECT System Fernando Pereira Instituto Superior Ttcnico - SecGSo de TelecomunicaG6es AV. Rovisco Pais, 1096 Lisboa Codex, PORTUGAL

[11] Vibrotactilefeedbacktoaidblindusersofmobileguides Giuseppe Ghiani,BarbaraLeporini,FabioPaterno` _ Istituto diScienzaeTecnologiedell'Informazione,ConsiglioNazionaledelleRicerche,ViaMoruzzi1,56124Pisa,Italy

[12] Chang, C. O'Sullivan. "Audio Haptic Feedback in Mobile Phones" in Proceedings of CHI 2005, April 2-7, Portland, Oregon, USA, pp. 1265–1267, ACM, April 2005.

[13] Toward mobile entertainment: A paradigm for narrative-based audio only games Timothy E. Rodena,_, Ian Parberryb, David Ducrestc a Center for Advanced Computer Studies, University of Louisiana at Lafayette, United States Received 1 October 2005; received in revised form 1 July 2006; accepted 1 July 2006 Available online 29 April 2007

[14] Mobile Phone Assistance for the Blinds.

[15] Kari Kallinen *, Niklas Ravaja "Emotion-related effects of speech rate and risingvs. falling background music melody during audio news: the moderating influence of personality Knowledge Media Laboratory, Center for Knowledge and Innovation Science Direct, Elsevere ,September 2003

[16] Olivier Collignona,b,□, Simon Girarda, Frederic Gosselina, Sylvain Roya, Dave Saint-Amoura,d, Maryse Lassondea,c, Franco Leporea,caCentre de Recherche en Neuropsychologie et Cognition (CERNEC " Audio-visual integration of emotion expression"), Université de Montréal, Montréal, Canada , Science Direct, Elsvere B R A I N R E S E A R C H 1 2 4 2 (2 0 0 8) 1 2 6 – 1 3 5

[17] SAMI-http://msdn.microsoft.com/en-us/library/ms971327.aspx

[18] Electronic Mail for the Blind User, Universiti Teknologi PETRONAS, Seri Iskandar, Final Year Project 2002

[19]. McDermott, J.H., Oxenham, J.A.,"Music perception, pitch, and the auditory system. Current Opinion in Neurobiology 18, 1–12 (2008)

[20]. Altinsoy, M.E.: The influence of frequency on the integration of auditory and tactile information. In: Proceedings of the 18th International Congress on Acoustics (ICA), Kyoto, Japan (2004)

[21] Occelli, V., Gillmeister, H., Forster, B., Spence, C., Zampini, M.: Unimodal and Crossmodal Audiotactile Frequency Matching in the Flutter Range. In: Altinsoy, M.E., Jekosch, U., Brewster, S. (eds.) HAID 2009. LNCS, vol. 5763. Springer, Heidelberg (2009)

[22] Dowling, J., Maeder, A., Boldes, W.: A PDA based artificial human vision simulator. In: Proceedings of the WDIC 2005, APRS Workshop on Digital Image Computing. Griffith University 2005, pp. 109–114 (2005)

[23] Sánchez, J., Aguayo, F.: Mobile Messenger for the Blind. In: Stephanidis, C., Pieper, M.(eds.) ERCIM Ws UI4ALL 2006. LNCS, vol. 4397, pp. 369–385. Springer, Heidelberg (2007)

[24] G. Chen, H. Tan, and X. Chen, "Audio Segmentation via the Similarity Measure of Audio Feature Vectors", Wuhan UniversityJournal of Natural Sciences, Vol. 10, No. 5, pp.833-837, 2005.

[25] C. Panagiotakis, and G. Tziritas, "A Speech/MusicDiscriminator Based on RMS and Zero-Crossings", IEEE Transactions on Multimedia, pp. 155-166, 2005.

[26] G. Tzanetakis, and P. Cook, "A Framework for Audio AnalysisBased on Classification and Temporal Segmentation",EUROMICRO, pp. 2061- 2067, 1999.

[27] M.S. Spina, and V.W. Zue, "Automatic Transcription of General Audio Data: Preliminary Analyses", pp. 594-597, ICSLP 96, 1996.

[28] H. Aronowitz, "Segmental Modeling for Audio Segmentation", Acoustics, Speech and Signal Processing, pp. 393-396, 2007.

[29] B. Ramabhadran, J. Huang, U. Chaudhari, G.Iyengar, and H.J. Nock, "Impact of Audio Segmentation and Segment Clustering on Automated Transcription Accuracy of Large Spoken Archives", Proc. EuroSpeech, pp. 2589-2592, 2003.

[30] S.S. Chen, and P.S. Gopalakrishnan, "Speaker Environment and Channel Change Detection and Clustering via The Bayesian Information Criterion", in *DARPA* speech recognition workshop, pp. 127-132, 1998.

[31] H. Meinedo, and J. Neto, "A Stream-based Audio Segmentation, Classification and Clustering Pre-processing System for Broadcast News using ANN Models", INTERSPEECH, pp.237-240, 2005

[32] Multigroup Classification of Audio Signals Using Time–Frequency Parameters

[33] J. Foote, "Content-based retrieval of music and audio," *Proc. SPIE, Multimedia Storage and Archiving Systems II*, pp. 138–147, 1997.

[34] E. Wold, T. Blum, D. Keislar, and J. Wheaton, "Content-based classification, search, and retrieval of audio," *IEEE Multimedia*, pp. 27–36, 1996.

[35] C. R. Fuller and A. H. von Flotow, "Active Control of Sound and Vibration", in IEEE Control System, Vol. 14, pp. 9-19, December 1995.

[36] Sorensen. H., Jones. D., Heideman. M., Burrus. C."Real-valued fast Fourier transform algorithms", in IEEETransactions on Signal Processing, Vol. 35, pp. 849-863,June 1987.

[37] A globally convergent frequency estimator, Liu Hsu Ortega, R. Damm, G.
Automatic Control, IEEE Transactions on
On page(s): 698 - 713 , Volume: 44 Issue: 4, Apr 1999

[38] http://new.jpost.com/Health/Article.aspx?id=211902

[39] Mr Moses, Malaysian Association for the Blind (MAB), Kuala Lumpur (interview)

[40] http://zone.ni.com/devzone/cda/tut/p/id/4278

[42] Dongju Chi, Yonghe You,"Sound specific vibration interface using digital signal processing" in IEEETransactions on Signal Processing,2008

[43] http://en.wikipedia.org/wiki/Band-stop_filter

APPENDIXES

Interview set questions.

Part (a); general questions

1. How is the level of confidence of blind people in doing something?(is it passive? Do they want to mix with others? What about in doing the same activities as of normal people?)
2. Is the brain processing of the blind is slower than of normal people in receiving information processing, imagine what actually happen during the football match?
3. How do they watch movies or watching television programs?
4. Can they understand on what actually happen during the playback/movies?
5. What is the different between blind since birth and blind caused by accident or disease (halfway) in imagine things? When receiving information? When understanding instructions?
6. Is sound really helpful in guiding the blind?
7. What kind of sound assistive technology that exist in NCBM?
8. By having sound effects/ auditory queue, is it helpful? If yes, explain how?

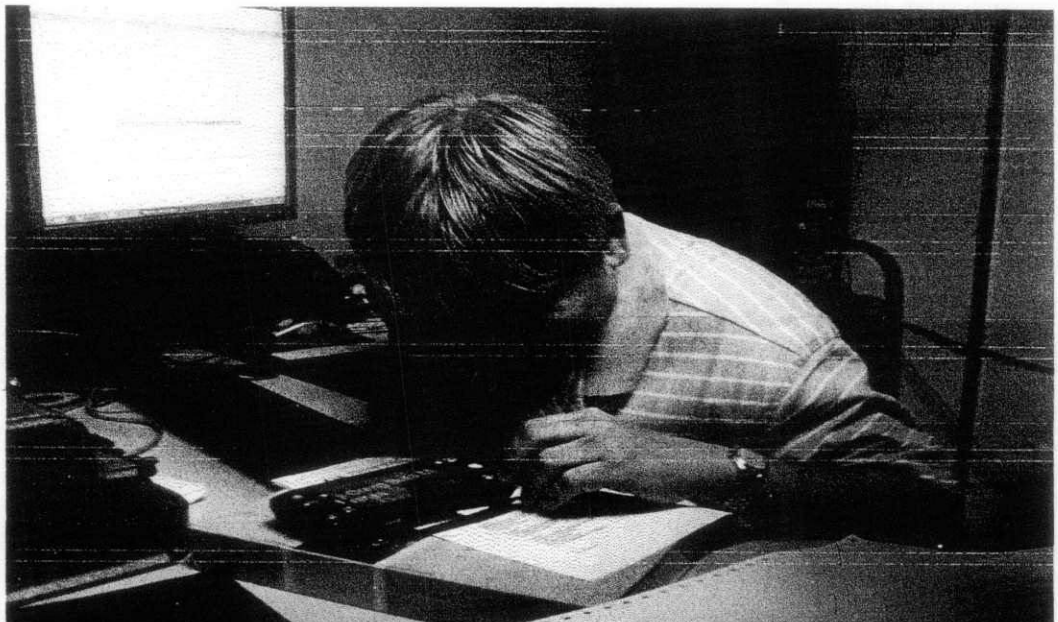
Part (b); football

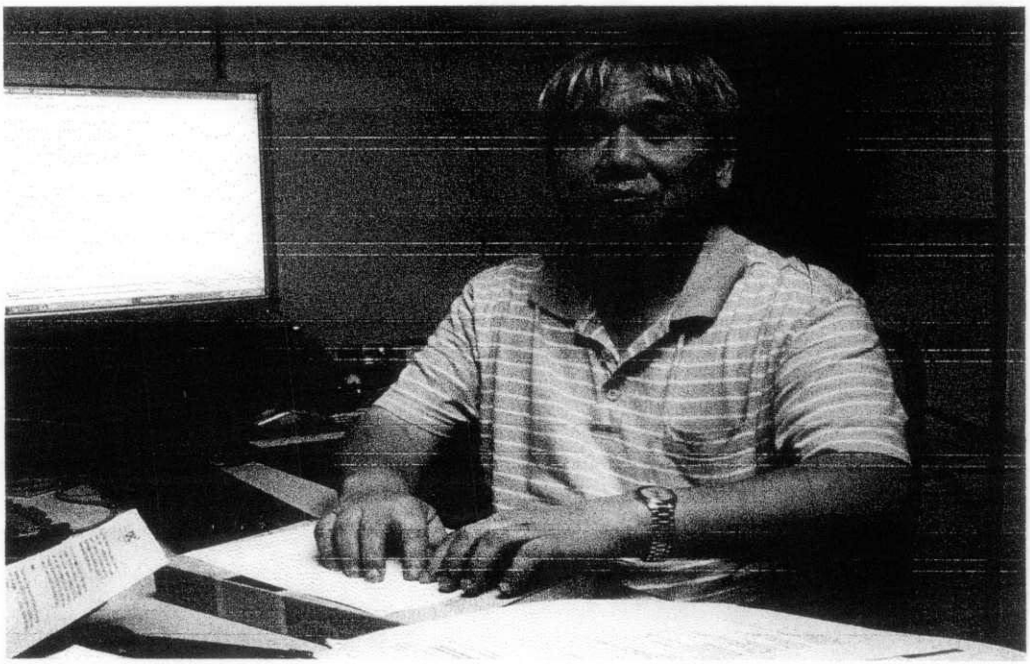
1. Do all blind people love sport?
2. Age of blind people who played football?
3. How do they get to know on how to play football?
4. How do they play football? Who assist them? How do they detect the ball? Use what kind of technology?

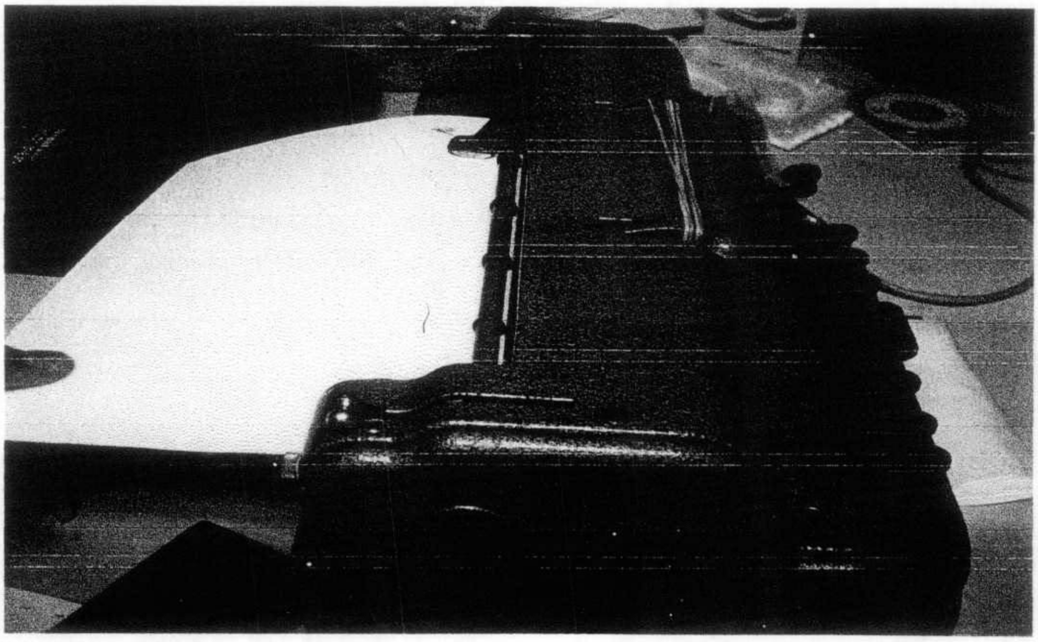
Part (c)

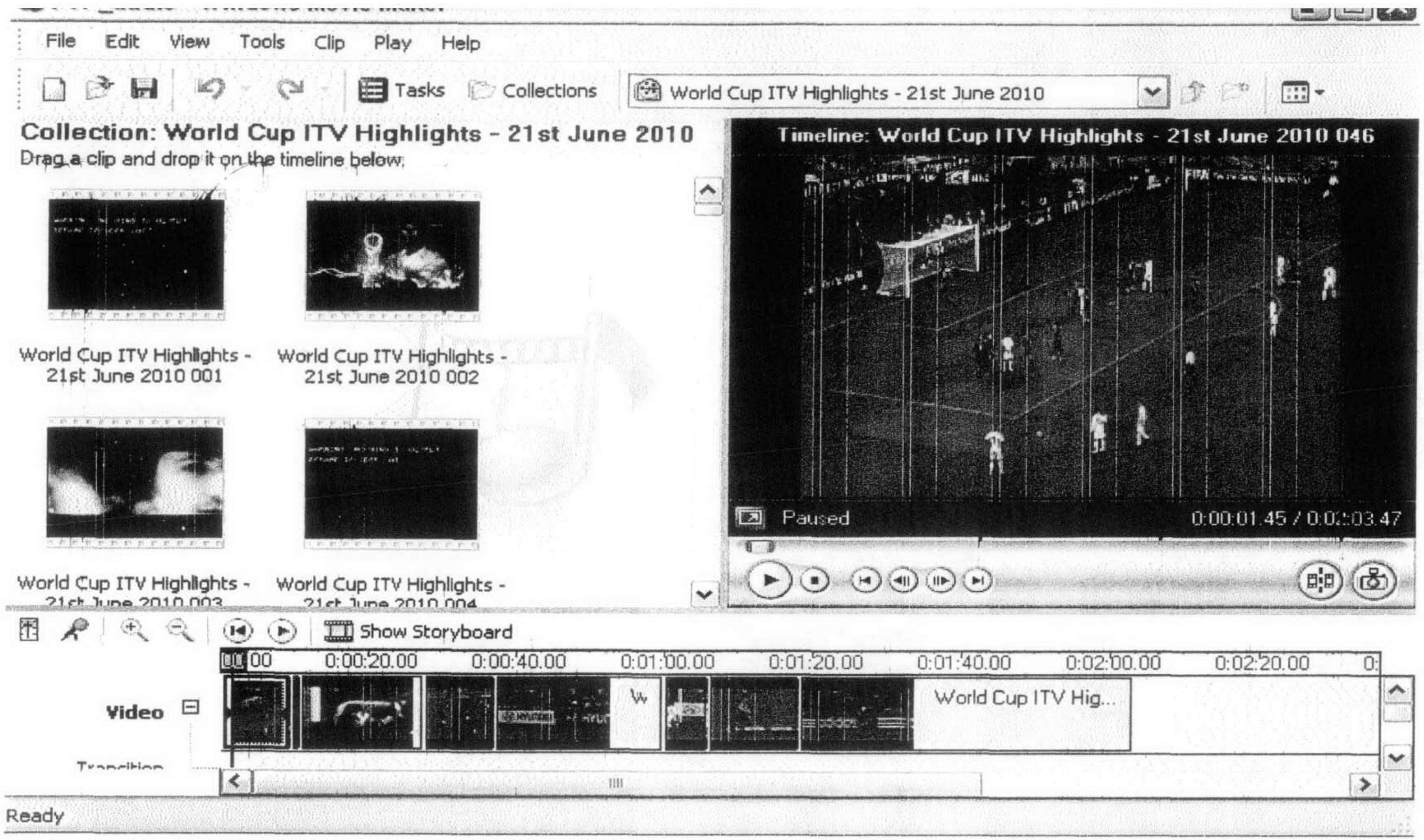
1. What do the blind people think on our effort to help them? And the effort did by government or any other blind association?
2. In terms of sound, how do they hope sound can help them?(enhancement on sound)
3. For the whole project, how do they hope we can help or enhance the technology?

Appendix: Figures show the pictures taken while conducting the interview at MAB.









Appendix: Figure shows the screenshot of Video Playback which need to be cut into shorter time for sampling

Appendix: Figures show the GANTT Chart for the project

Gannt Chart for Final Year Project (January/Mei 2011)

No	Task	Duration	Start	Finish	Mo							
		Days			1	2	3	4	5	6	7	8
1	Immersive Assisting Tools and Application for the Visually Impaired People (Audi		24/01/2011	26/08/2011								
	FYP 1		24/01/2011	25/04/2011								
2	Stage 1 - Proposal Approval	12	02/02/2011	14/0/2011								
3	Endorsement on the project topic	1	02/02/2011	02/02/2011								
4	Project proposal submission	1	07/02/2011	07/02/2011								
5	Stage 2 - Research and Development											
6	Phase 1 - Critical review of related works	47	07/02/2011	25/03/2011								
7	Conduct research on accessible techniques	28	07/02/2011	04/03/2011								
8	Submission of Extended Proposal	1	04/03/2011	04/03/2011								
9	analysis and design on Audio	21	04/03/2011	25/03/2011								
10	analysis and design of Audio Frequency Segmentation	21	04/03/2011	25/03/2011								
11	proposal defense and progress evaluation	1	25/03/2011	25/03/2011								
12	Phase2 - simulation development	30	25/03/2011	25/04/2011								
13	develop a preliminary framework on Audio Frequency Segment	7	25/03/2011	01/04/2011								
14	simulation development of Audio Extraction + Frequency Segm	7	01/04/2011	08/04/2011								
15	submission of interim report	1	08/04/2011	08/04/2011								
16	testing and evaluation of the simulation	25	08/04/2011	25/04/2011								
17	Submission of technical Report	1	25/04/2011	25/04/2011								

	FYP 2		25/04/2011	26/08/2011								
18	phase 3 - test-bed development											
19	development of Audio Frequency + Stored Sound Matching											
20	testing and evaluation of the test-bed											
21	Submission of Progress Report											
22	phase 4 - application prototype and development											
23	development of the system prototype											
24	demo of the system prototype											
25	Evaluate results											
26	Submission of Progress Report											
27	Produce Dissertation											
28	Stage 3 - Documents and Publications											
29	Submission of project Dissertation											