CERTIFICATION OF APPROVAL

Search Engine for WAV files

by

Noor Ida Rahayu Bt. Ahmad

A project dissertation submitted to the
Information System Programme
Universiti Teknologi PETRONAS
in partial fulfillment of the requirement for the
BACHELOR OF TECHNOLOGY (Hons)
(INFORMATION SYSTEM)

Approved by,

(Dr. Abas Md. Said)

UNIVERSITI TEKNOLOGI PETRONAS
TRONOH, PERAK
June 2004

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 except as specified in the references and acknowledgements, and that the original work contained herein have not been undertaken or done by unspecified sources or persons.

NOOR IDA RAHAYU BT. AHMAD

ABSTRACT

This project deals with matter pertaining to the application of search engine for WAV files. The objective of this project is to create an application that is capable of searching for WAV files by comparing the criteria of the input and the WAV files in the database. The emphasis is placed on the aims at providing user with the technology that will help them to analyze while comparing the differences between the input given and the original WAV file in the database. This application is meant to assist user that have problem searching for the matching title of a particular WAV files. The main features of the application are that 1) it receives any kind of input form from the microphone, 2) it then compares and matches the criteria of WAV files from the database using specified method. The scope of study is the process of how audio recognition work and how it could be compared based on certain identified criteria to find the matches. It basically focuses and emphasizes on WAV file format comparison but in a broader perspective of research in WAV file criteria and attributes. I don't have any specified framework in doing my project, yet I follow iterative guidelines where much of the processes are done in parallel and without planning. The finding will be focusing on the issues on how to capture, compare and match WAV files format. Basically, by using the underlying concepts the project will be able to compare the input WAV format and match with the original intended song easily and faster.

ACKNOWLEDGEMENT

Firstly, I would like to express deepest gratitude to Allah S.W.T for giving me the strength, wisdom, and patient in order to complete this project within the time allocated.

I would also like to thanks to those people that have helped me to complete this project. First and foremost, deepest gratitude goes to my Final Year Project Supervisor, Dr. Abas Md. Said for the continuous support, guidance, and motivation which have aid and give me the inspiration and strength to undergo this project successfully. Without your assistance and concern, I might not have completed it.

Not forgotten a special thanks and love to my beloved parents Encik Ahmad B. Saleh and Puan Normah Bt. Ali and also my family members for all the morale support, love, and encouragement for me to complete the project.

A special thank go to Mr. Jafreezal Jefree, Mr Jale, Ms Kavita, lecturers and tutors in Universiti Teknologi PETRONAS whom had helped me in giving ideas and advises especially in developing the prototype.

Last but not least, to all my colleagues in Universiti Teknologi PETRONAS, thank you for being there when I need your help throughout the project development. I really appreciate your help especially for sharing experiences, knowledge and contributing brilliant ideas that are much needed in my project.

TABLE OF CONTENTS

CERTIFICATION OF APPROVAL i

CERTIFICATION OF ORIGINALITY ii

ABSTRACT iii

ACKNOWLEDGEMENT iv

TABLE OF CONTENT v

LIST OF FIGURES vi

CHAPTER 1 INTRODUCTION

1.1 Background of Study. 1

1.2 Problem Statement 3

1.3 Objectives and Scope of study. 4

CHAPTER 2 LITERATURE REVIEW AND THEORY

2.1 Nature of Sound. 6

2.2 WAV file format. 7

2.3 Audio search engine 8

2.4 Digital Signal Processing 11

2.5 Nature and content of signals 13

CHAPTER 3 METHODOLOGY AND PROJECT WORK

3.1 Project Work 16

3.2 Project Work Framework. 20

3.3 Tools required 21

CHAPTER 4 RESULTS AND DISCUSSION

4.1 Results. 22

4.2 Discussion. 34

CHAPTER 5 CONCLUSION

5.1 Summary of the project 36

5.2 Future recommendation 37

5.3 Conclusion 39

REFERENCES 40

APPENDICES

Appendix 1: Use Case Diagram Compare WAV files

Appendix 2: Use Case Diagram Extracting Matched WAV files

Appendix 3: MATLAB Programming Code (M-file)

Appendix 4: Final Presentation Slide

Appendix 5: Advertisement on Digi

Appendix 6: Compilation of FYP CD

LIST OF FIGURES

Figure 1.1	The transmission of sound through an electrical system
Figure 2.1	Construction of Sound
Figure 3.1	Project Work Framework
Figure 4.1	Example of wave sign from the original WAV file in the database
Figure 4.2	Example of wave sign from the input human voice
Figure 4.3	Example of wave sign after alteration are made
Figure 4.4	Simulink Model
Figure 4.5	Graphical User Interface
Figure 4.6	Error Graph Plotting (yout)
Figure 4.7	Error Graph Plotting (y1out)
Figure 4.8	Error Graph Plotting (yout0)
Figure 4.9	Result Display

CHAPTER 1
INTRODUCTION

Chapter 1 consists the basic information of the project, comprises of its background, its problem statement, its objectives and the scope involved. This section also described the intention of capturing, comparing and finding the method to extract audio databases specifically the WAV file format and to be able to tell and visualize the differences between both files.

1.1 BACKGROUND OF STUDY

Ross Kirk et al (1999, p.4) (1) points out that "In this age of rapid development of the new high-technology era, it is again our own concern to keep a sensible perspective by reminding ourselves that music has been the fundamental part of every human society for thousand of years. " Meaning that the technological advances have allowed large proportions of the population that to music and computing facilities as an integral part of everyday life.

Audio Search Engine is the concept whereby it is an application that searches audio files through different type of input such as voice humming, recorded instrumental or original recorded song and returns the comparison results of both WAV files. The WAV file format itself is the mean of recording an actual sound at a high-frequency digital sampling (accurate technique) and can thereafter faithfully reproduce that sound repeatedly. This sound file format is often large in size to facilitate the storage of all those data points required to reproduce the sound without any depreciation of the sound quality. In the world of music, analog sound refers to a nondigital tape or vinyl record recording of sound. It is also the name for an electronic signal that carries its information or sound as a continuously fluctuating voltage value. When the sound is not in digital form and yet is stored in analog form, we need to carry the digitizing process. Digitizing is the process of converting an analog signal to a digital one, in order to do this, we need analog-to-digital converter that is used to recording sound on the computer through a microphone, and this is actually using the A-to-D converter of the computer's sound interface.

Audio and music technology are concerned with the recording, transmission and manipulation of the acoustic signal within electrical systems. The sound pressure waves falls on a moveable diaphragm in a microphone and this causes the diaphragm to move in direct sympathy with the sound waves. Figure 1.1 outlines the fundamental understanding of the nature of signals (the conceptualization of signal waves). This electrical analogue are then can be manipulated and stored. Any changes which been make upon it will have an equivalent effect on its analogue in the sound world.

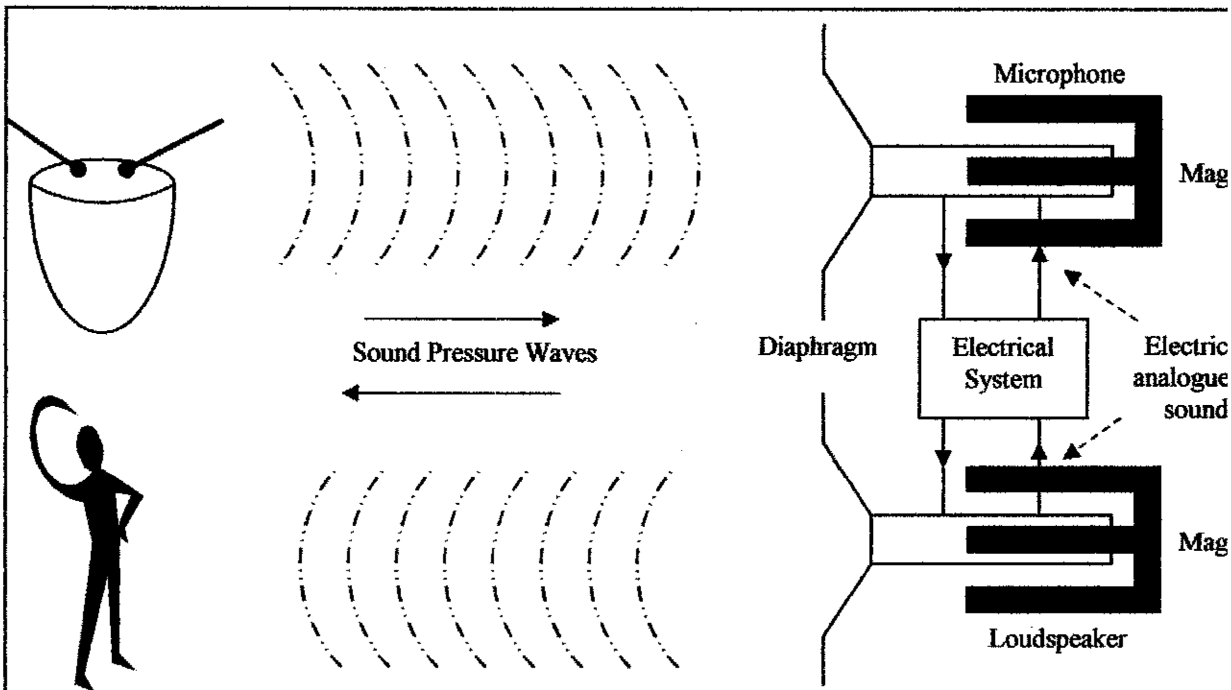


Figure 1.1 The transmission of sound through an electrical system

Source: Ross Kirk et al (1999, p.42) (1)

1.2 PROBLEM STATEMENT

1.2.1 Problem Identification

Music is a routine for every human being. Day by day they listen to music, whether in car, at home, at the office or even shopping. Everywhere they go there will be music around. The main problem is, they might seem to get the lyrics correct but most of them don't exactly know the title of the song.

Here, the issue has been identified and recognized. In order to help solve it, this WAV files search engine application would be a bonus to everyone. Most of the search engine available on the World Wide Web searches for text files not WAV files. This application only search for song title by comparing the signal of the input and the signal in user own hard disk database.

1.2.2 Significant of the project

Implementation of this project would mean that user will have the advantage of searching for WAV files through different types of input direct from the hard disk database. User will no more find difficulties in searching for a matching song titles. All they need is just input device that will capture their voice and searches for the matching WAV files in the computer database.

1.3 OBJECTIVES and SCOPE of STUDY

The project concerns with the application of audio search engine through comparison of input WAV files and original hard disk database stored in WAV files format. The main purpose of this application is to find differentiation and accuracy while visualizing different type of WAV file input with the original one.

The main features of the application are that 1) it receives any kind of input form, 2) it then compares and matches the criteria of WAV files from the hard disk using specified software and method and 3) do the matching while displaying the reasons for differences occur. This project is meant to extract some of audio databases using specified method. It also intended to help visualize the different criteria of multiple WAV files through comparison process.

The objectives are as follows:

1. To do research on the audio aspect of a WAV files
2. To do research on how the to capture different kind of input with the most minimal noise rate
3. To try extracting some of audio databases using specified method
4. To show the different outcome of WAV files after comparing being made
5. To understand the underlying concepts of search engine for WAV files
6. To understand why WAV file can't be accurately matched

The background of study is basically a circulation of how the capturing of different input forms can resulted in different outcome after comparison of WAV file are made.

The scopes of study are as follows:

1. Focuses on the concepts of capturing different input forms which then resulted in different outcome after comparison are made
2. Research on ways to extracting some databases using specified method
3. Concentrate on the criteria that differentiate wave signal
4. In-depth coverage of WAV files format

1.3.1 The Relevancy of the Project

The purpose of the system is to provide user with the technology that will help them to search for WAV files that match the input given through the input device through comparison. Its objective is to try extracting some of the audio databases using some specified method. It also show the different outcome of WAV files after comparing being made while understand the real concept on why WAV file can't be 100% accurately matched. This application can only be implemented in a Personal Computer. It only searches for matching WAV files in the hard disk database. Maybe later it would be enhanced to be implemented to other platform.

1.3.2 Feasibility of the Project within the Scope and Time frame

A time frame given is about four months in order to complete this project. In order to achieve the objective of this project, I have planned everything in accordance to its aim based on the allocated time frame. Eventhough, there is no specific task allocation in this project, I am sure that all the necessary planning that are done in parallel and iteratively are really good enough to complete the project.

medium it travels including air, water, gold, and granite, and 4) Human Ear that serves as an astounding transducer, converting sound energy to mechanical energy to a nerve impulse which is transmitted to the brain.(2)

It is also being stated in a reference textbook called The Physics Hypertextbook™, © 1998-2003 by Glenn Elert that sound is a longitudinal, mechanical wave, which mean sound can travel through any medium, but it cannot travel through a vacuum and there is no sound in outer space.

Sound is a variation in pressure. A region of increased pressure on a sound wave is called a compression (or condensation). A region of decreased pressure on a sound wave is called a rarefaction (or dilation).

2.2 WAV file format

So, basically sound can come with different file extension. In this project scope of study it only involves .WAV files. WAV is Microsoft's audio format. Since Windows 3.1, WAV has been the native format for sound within the Windows environment. Needless to say, this makes it one of the most common sound formats on the Web. Most browsers support the .wav format with their internal sound players. A .wav-formatted file has this extension for example Test123.wav.

WAV files are probably the simplest of the common formats for storing audio samples. It is compressed although might consume large capacity of storage. WAV file store samples "in the raw" (from the aspect of sound construction) where no pre-processing is required other that formatting of the data. From the literature review of WAV itself, it consists of three "chunks" of information: The RIFF chunk which identifies the file as a WAV file, The FORMAT chunk which identifies parameters such as sample rate and the DATA chunk which contains the actual data (samples).

2.3 Audio Search Engine

We can see that the progress of search engine development are widely focuses on keywords or text, but the development of WAV file search engine intentionally to search files by input either voice humming, recorded instrument, or recorded original song is still new. But there are several researches have been done all over the world on this project. There are more a less the same with what I'm researching on, but the approach and method used might somehow differ in a way, and yet still can be used as a guide for my project. The following paragraph will actually discuss on some of the project research that are much related to what I am doing.

Ghias et al. (3) published one of the early papers on query by singing. They applied autocorrelation of range errors to obtain the fundamental frequency, and the pitch vector is then cut into notes (portion by portion). To accommodate the problem of different staring key, the obtained notes that are converted to ternary contour of three characters: U (up, meaning this note is higher than the previous one), R (repeat, meaning this note is the same as the previous one), D (down, meaning this note is lower than the previous one). The comparison engine, based on the longest common subsequence, is then applied to the ternary contours to find the most likely song. However, due to the limited computing power at that time, their pitch tracking takes 20-45 seconds, and there were only 183 songs in the database.

R. J. McNab et al. (4), in collaboration with New Zealand Digital Library, have published several papers on their experiment of query by singing. They applied Gold-Rabiner algorithm for pitch tracking, and the pitch vector was then cut into notes based on energy levels and transition amounts. As for my project, energy levels and transition amounts referred to the length and frequency of amplitude. There were about 9400 songs in their database and they are the first one to put their system on the web. Their system, though lack of performance evaluation in terms of recognition rate, is still considered an excellent example of content-based music retrieval for real-world applications.

A research project published by Jyh-Shing et al. (5) on "content-based music retrieval via acoustic input" are another system that a content-based music retrieval system that can take a user's acoustic input (S-second clip of singing or humming) via a microphone and then retrieve the intended song from a database containing over 3000 candidate songs. The system, known as Super MBox, demonstrates the feasibility of real-time music retrieval with a high success rate. Super MBox first takes the user's acoustic input from a microphone and converts it into a pitch vector.

Then a hierarchical filtering method is used to first filter out 80% unlikely candidates and then compares the query input with the remaining 20% candidates in a detailed manner. The output of Super MBox is a ranked song list according to the computed similarity scores. A brief mathematical analysis of the two-step HFM is given to explain how to derive the optimum parameters of the comparison engine. The proposed HFM and its analysis framework can be directly applied to other multimedia information retrieval systems.

According to a research done by John H.L. Hansen et al. (6) on "Transcript-Free Search of Audio Archives for the National Gallery of the Spoken Word", although the problem of audio stream search is relatively new, it is related to a number of previous research problems. Many systems developed for audio search, however, assume the existence of associated text or a clean audio stream (less noise rate). Direct information retrieval via audio mining generally focuses of relatively noise-free, single-speaker recordings. Alternative methods have included ways to time-compress or modify speech to allow listeners the ability to skim more quickly through recorded audio data. While keyword spotting system can generally be used for topic or phrase search applications, the system must be able to recover from errors in both a user's query and in rank-ordered phrase hypotheses in the stream. Phrase search focuses more on locating a single requested occurrence, whereas keyword or topic spotting systems assume a number of possible outcomes.

Their project are much related and similar to my project but the approach they take is rather too way complicated and complex. Much of it involves mathematical computation and arithmetic calculation which is not 100% applicable to my project. My approach is rather easy and involves simple rules of calculation. Cutting the pitch vector into notes which then they translate it to ternary contour of three characters is also their one approach of doing it. Both of the music retrieval systems published by Ghias et al. and McNab et al. are designed to extract music notes from the identified pitch vectors, and then comparison procedures based on the extracted notes are invoked to compute similarity scores. However, note extraction (or segmentation) is error prone due to the continuity in pitch vectors obtained from human's acoustic input. So do what John H.L. Hansen et al. have done, it considers the noise occurrence and able to skim down the recorded audio data so that the quality of the input is better to be compared.

2.4 Digital Signal Processing

Signal processing is the process of extracting relevant information from the speech signal in an efficient, robust manner, meaning extracting the whole composition of the sound e.g. pitch, frequency, amplitude. Included in a signal processing is the form of spectral analysis used to characterize the time-varying properties of the speech signal as well as various types of signal processing (and postprocessing) to make the speech signal robust to the recording environment (signal enhancement). In this research computing environment, potential algorithms are first tested on small amounts of data (portion by portion testing algorithms) which have been chosen or simulated to mimic certain important situations. Less length of the song, less error will be incurred. Later, ever-increasing amounts of realistic data are used to refine and test the algorithm. At some point, sufficient performance is achieved on adequate amounts of representative data, and the algorithm is declared ready for use.

In this project it is more focus on the real time concept of digital signal processing. Referring to reference textbook of "Real-Time Signal Processing" (1999, pg 1) (7) it is explained that, in a real time signal processing application, the samples of the digital signals arrive at the input of the processing system. This means that, these samples will arrive over some interval of time, until a group of samples known as a frame, have arrived then only it is sufficient to begin processing. Digital signal processing applications seek to extract useful information from an incoming set of signals, represented by sequences of numbers. The algorithms or computational prescriptions, for extracting this information are originally invented and developed in general-purpose computing environments in which the speed of modifying and retesting the algorithm is more important than the speed of executing the algorithm. It is because modifying and retesting the algorithm is used to get the best algorithm rather than executing the algorithm which might not guarantee the best even if it runs without errors.

There are few categories of real time signal processing systems on the basis of the length of the range of input samples and the sequence of computation on the input samples that produces the output samples. The first one is stream processing whereby, it proceeds on a sample-by-sample basis upon the input sample sequence, performing an identical set of computations upon each sample of the analysis frame and completing these computations before the next sample arrives. The second one is block processing, which is a technique to store incoming samples in group as they arrive. After the arrival of a sufficient number of samples as set by the nature of the implementation, processing begins on the group of samples. Because all input samples of the group, or block are available simultaneously, processing can access the samples in a random manner rather than being restricted to sequential access. Last but not least is the vector processing. This is the category where, it applies to processes that operate on many signals at once which are combined as a vector. If these samples derive from the same signal, the situation of block processing will be obtained. Vector processing consists of operating o multiple simultaneously-arriving samples from several signals. As for a single signal channel, this processing may either occur in a stream mode or a block mode. (1999, pg 14) (7).

To cast signal comparison algorithms into real time, both algorithmic innovation (the mathematical error range calculation) and architectural structural design (the usage of simulink model, block parameter) are important to speeding computation. Such innovations include limiting the search space and using sequential decision-making or coarse-to-fine search strategies.

Algorithms for signal comparison are distinguished by several features that affect the real-time implementation. First, it consists of not only a matching mode that compares an unknown input to a collection of previously stored data or models, but also training mode by which patterns are built from models or data and compared to signals to affect the step of recognition. The training mode itself uses the matching mode to guide the training. Second, the pattern comparison operation consists of a hierarchy of levels. At the lowest level, a pattern's smallest units are compared to those of another pattern. At the next higher level, these local distances are summed along consecutive units according to various mapping paths to form global or accumulated distances. This accumulation allows the formation of words from phones or frames, or objects form pixels or features. Third, during this hierarchical combination, several paths that map local distances of one pattern to those of another are analyzed, seeking to compensate for distortion by the choice of path used in the accumulation of the final global distance.

2.5 Nature and content of signals

Understanding the nature of the electrical analogue of sound signals require us to first visualize the construction of sound. Figure 2.1 visualize the construction of sound. An oscilloscope used to visualize the output of a microphone while giving a closer view of the oscilloscope screen for the words digital audio spoken into the microphone are shown below.

An oscilloscope is an electrical instrument which displays variations in voltage (vertical axis) against time (horizontal axis). It therefore traces out fluctuations in voltage over time on its screen. Voltage measure of the electrical 'pressure' which forces current along the wires of a circuit, and so it is another analogue of sound in our microphone example. The figures shows the waveform which might be observed on an oscilloscope for the words 'digital audio' spoken into the microphone. The output therefore might seem accurate but, it would be probably hard-pressed to classify it in a meaningful form, except to describe it as a wiggly line. This indicates that it is more preferable to consider a simpler signal than the sound of spoken phrase so that it is more understandable and also easy to develop in a simpler model of analyzing the content of a signal.

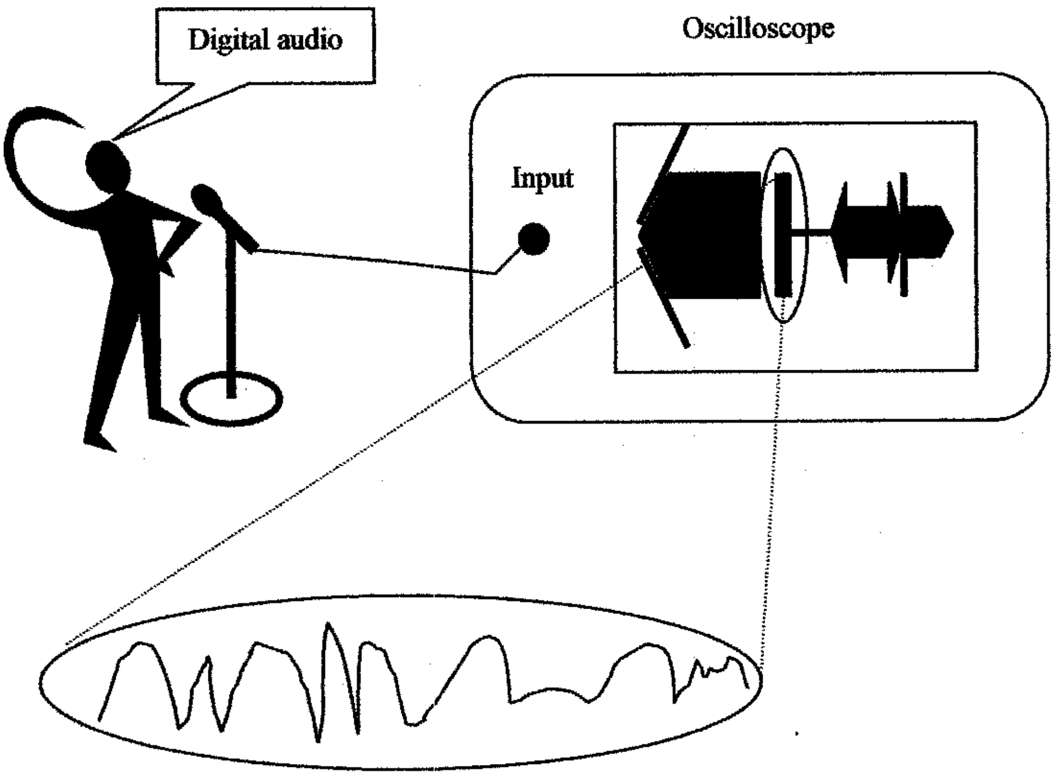


Figure 2.1 Construction of Sound

Source: Ross Kirk et al (1999, p.44) (1)

To attain real time implementation, the pattern search is often speeded by limiting the search space. In many cases in the project, limiting the search space degrades performance, although the intent and practice is generally to obtain factors of two to five speedups with no more than a 10 percent increase in error rate.

These are actually few examples of real research project being done all over the world. Most of it are good in certain ways and can be granted as an example and guideline to what my project is. What actually differs from each other is just the way they operate the input data to achieve the aim. In my concerns, the aim is to make comparison on why different inputs have different outcomes when comparing to its original WAV file. Furthermore, intentionally I am trying to extract some of audio databases using different methods. It however provides a good exposure for my project as it can help to educate and train me in doing a more detailed research on the project itself. Their examples contribute a lot to my research giving an extra credit for my project development as they provide the tools and raw information that are sufficient and beneficial enough to achieve my goals and objectives. The way they perceive and solve their issues arising in their project have taught me how to handle my issues arising too.

CHAPTER 3
METHODOLOGY and PROJECT WORK

Chapter 3 consists information on the methodology adopted and the flow of overall project work. This section also described the concepts of the methodology and the relevancy of the methodology to project work. Till the date of this report (9^{th} April 2004), I've already managed to complete almost 90% of the phases involved. Tools used throughout this project will also be discussed and explain in this section.

3.1 Project Work

Throughout the development of the application, there is no specific framework for describing the phases involved in developing this project. I don't practice a specific methodology which specifies which phase comes first and which comes next. Overall the project is actually done in parallel and without any indicated phases, but the concept of the project development flow is actually visible according to my progress. First, when me and my supervisor discussed and have agreed on doing the project, I started to set the objectives, goals and scope. This is the process to ensure that the project is not out of the topic agreed upon earlier. When I first set the objectives and analyze appropriate scope, I did few research and studies related to the project. Sometimes, there is a need for scope redefine because it might change depending on the progress of the project.

Then, detailed discussion was held based on weekly basis to understand what really is needed to be done. I also need to educate myself and understand the arising issues of the project. As advised by my supervisor, I do research from time to time which mean that I do not wait until the due date of my thesis to search for information. It is a looping process as it will be repetitively done when there is a need for information seeking. I read a lot of magazines, newspapers articles, journals, and related reference books from the resource center to educate myself in order to understand each and every inch of the element in my project.

I then have to produce a Gantt chart just to be regard as a rough guideline of my project flow. Even though I don't really follow exactly the date stated there, but the Gantt chart provide me with the sufficient information on how I can achieve my objectives and eliminate the issues of time constraints. After related studies are being done, few reports need to be done and submitted to be evaluated by my supervisor. In addition to strengthen the research made, I have make appropriate finding on how the nature of sound, analyze the WAV file format and how to manipulate it. I also have the chance to learn software for editing sound files such as Sound Forge. Sound Forge supply the basic functions that help user to manipulate the WAV file in certain ways. Using this software, I am able to manipulate the file by altering the pitch, frequency, tone, speed, amplitude and the pattern of the sign. But as time passes and new things discovered, some of the functions and tools provided by the Sound Forge resulted in insufficient to manipulate and compare WAV files, so I research for other software that can actually help me to achieve my aim.

Gathering data and information on basic of sound files, understand the WAV format, distinguish it and manipulate it is done continuously and in parallel with other task that I am doing such as while I'm comparing the wav sign or learning new software that might be appropriate for my project. I have to get equipped with the basic nature of sounds files so that later I am able to manipulate and differentiate different factors of WAV format files and able to handle anything that concern to it.

Next, I sketched models and project flow that represent all the project processes, outputs and inputs involved in executing it, such as the conceptual and functional modeling of how the system flows and works. Progress report submitted is one of the outcomes of the model and project sketching which is a continuous, iterative process that allows me to understand, modifies, and eventually approves a working model of the project that meets the project goals, objectives, and scope. I also interact a lot with my friends to get their ideas especially what and how the project should works and functions. Refer to Appendix 1 for full sketches of the project flows and functions. The end result corresponds to the project scope and is flexible against changes because as fast as it is developed it can still be changed.

The most important aim in my project is to get the two different inputs of WAV file format criteria compared and finds the most matching one by 1.) Input e.g. voice, instrument, and exact song compared with the 2.) WAV file format in the database e.g. the exact song and 3.) finds the most matching file. End result will help me visualize and make instance corrections to the project I am developing. This will also allow details to be adjusted if necessary for example if the WAV sign need to be compressed or stretched to compare and matched with the original WAV file format in the database.

Comparing WAV file formats is the most time consuming task, to get the two different input matches, the display output in the screen need to be manipulated, there need to be a consequences whereby, it should consider the acceptable error rate range in the function. Since this project is still new and in the research progress it can be considered as challenging and need a lot of work and study. Therefore, up until now I still make research and studies on what other people have done that can be considered a little similar to what I am doing. The findings have helped me a lot in giving me ideas and also aid me in visualizing how I can achieve my project aim. Now, I also consult my supervisor and few lecturers that can give me ideas on how to solve the problem. I also learn about the Matlab as it will be the main tools to achieve my project. In my project development, I allow iteration whereby it helps to guarantee effectiveness and self-correction. I found out that in order to search for a WAV files by comparing the signal of amplitude, the pitch and speed is going to be complicated and rather time consuming but MAtlab have the alternative to the issues, it actually use digital signal processing of signals initially in the form of an analog electrical voltage or current, produced for example by a microphone or some other type of transducer.

There is no exactly a reason why I didn't have a list of phases as my methodology like anyone else perceived in their project development. Since this is just the development of a prototype and yet not the full system development like others in the research have done, so it is more wise if I do it on the basis of time constraints and scope of the project. I don't want to spend time planning and choosing a specified methodology for my project as the development and task allocations is done impromptu and without proper guided methodology. In the real life of completing my project, I do whatever come first and sometimes I did go back to what I have done before because there might need a changes in the task. Everything seems to be done simultaneously to ensure that time is obeyed and objectives are achieved.

Studies have shown that human beings almost never perform a complex task correctly the first time. However, people are extremely good at making an adequate beginning and then making many small refinements and improvements. So, I foresee and accept my problem in developing the project as something that encouraging and I fight to solve it rather than giving up. In this project I hold the principle of "application is always complete but never finished". Now, I am still polishing the project that I am developing to ensure that it is more visible to see and efficient to be used.

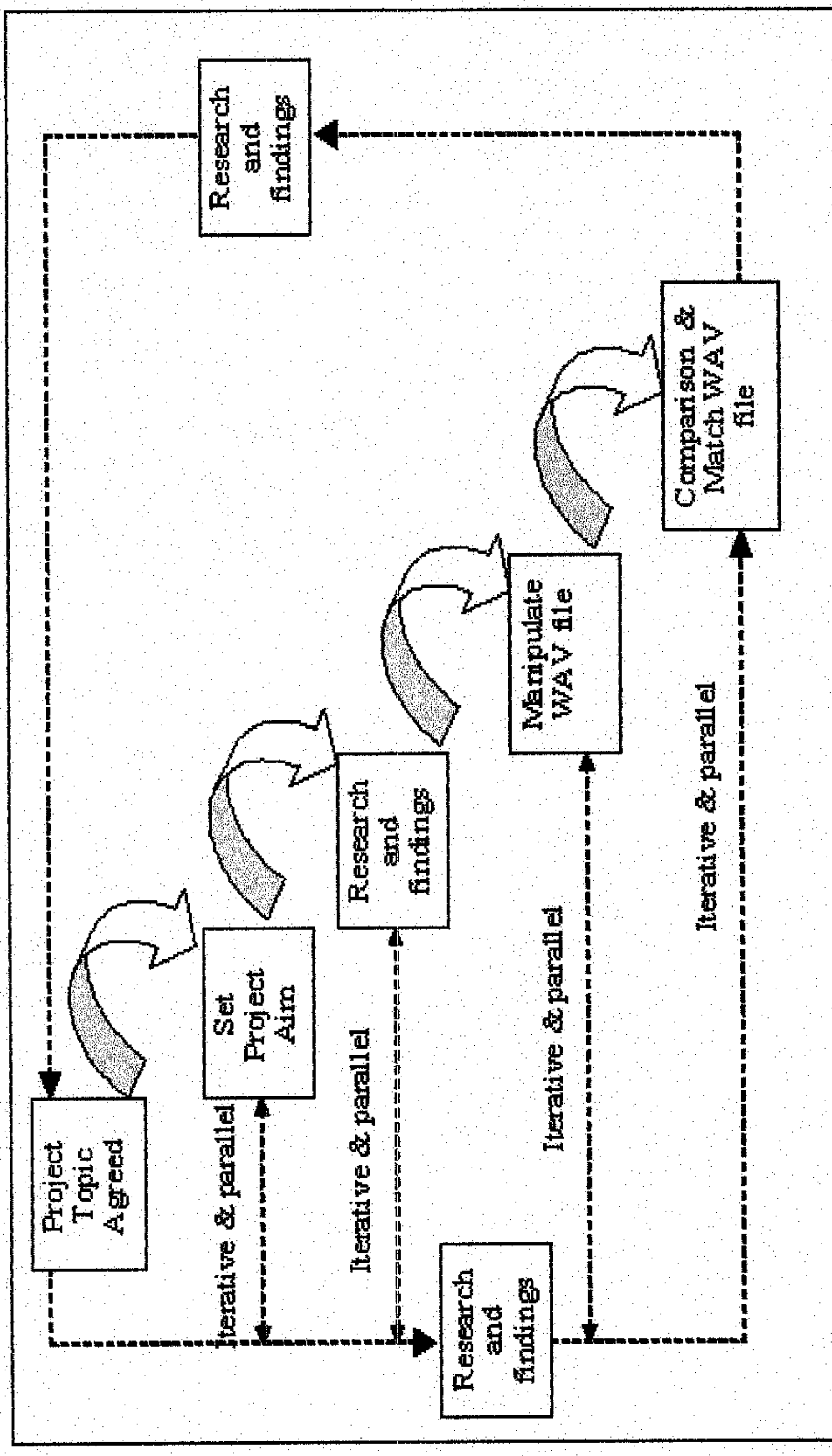


Figure 3.1 Project Work Framework

3.2 Tools Required

As stated in previous report, the tools listed are not final yet and any slight changes of tools usage might happen. After proper analysis and study, the most appropriate tool that will satisfy the need of my project is Matlab 6.1. It can be used to output the result of a given input, make appropriate comparison, and match it with the WAV file in the database.

The hardware will be:

- A PentiumII-equivalent or later processor at 233 MHz with 256 MB of RAM is recommended with Windows Operating System
- A microphone or some other sound input device to receive the sound is required. The microphone should be a high quality device with noise filters built in. The input capture rate will be significantly lower or perhaps even unacceptable with a poor microphone
- A recorder of any kind that is used to record original song and playback it into the microphone as input

CHAPTER 4
RESULTS and DISCUSSION

Chapter 4 compiles the current findings of the project work. There have been several interesting information, coming from journals and online resources. The results of the findings that have been decided are discussed in this chapter. What I theoretically understand from the research and study is then being conducted practically through the development of the prototype. How it actually can influence and aid my own research should be clearly outlined in the results and discussion chapter.

4.1 RESULTS

From all the findings and studies being made based on my project, I finally understand the fundamental concept of WAV file format. Currently, I am able to manipulate multiple different input forms of WAV files by stretching and compressing the signal and visualize the element of differentiation being made. What actually differ from those files is visually visible after proper alteration has been made to it such as adding noise and etc. The result of the comparison being made after the comparison have proved to me there are few criteria such as noise rate, accuracy of input rate, and error rate of recording should be considered as part of the digital signal processing of audio.

After deeper study and research on the project, I have learned a lot of new things which have helped me in achieving my project objectives. At the earlier stage of the project development, I have few issues that seem complicated and these issues have lead to project delay. The most complicated issue that arises is finding the way to manipulate WAV files. First, I have found a software that are usually used to record and manipulate sound called SoundForge.

Using SoundForge, I have the ability to record input WAV files and view the wave sign with the amplitude. Then, I can also open the original WAV file in the database itself and view the wave sign too. But, the limitation is that it do not supply user with the function tool to compare both WAV files although there are tools that able the user to add noise, compressed or stretched the wave signal and etc. I was wrong back then because, I have taken to much time on learning the software and yet the software doesn't really help me in achieving my objectives. It is good software even though not so appropriate and applicable for my project.

As we can see below using the SoundForge software, there are three different figures of wave sign from the same WAV file song. The first one is the original WAV file, the second is the same file but from human voice input, and last one is the same file but after modification have been made. Roughly, we can visualize that the wave sign are all in the same manner. What actually make these three figures different are the differences in the pitch blend, the high and low of the amplitude, and etc. Hence, the concept is still there. Figure 4.3 is a wave sign that is after modification being made, which mean that I have add the noise rate, blend the pitch and stretch the wave sign. But, the problem is that, SoundForge has no tools that are able to compare it and to do the matching.

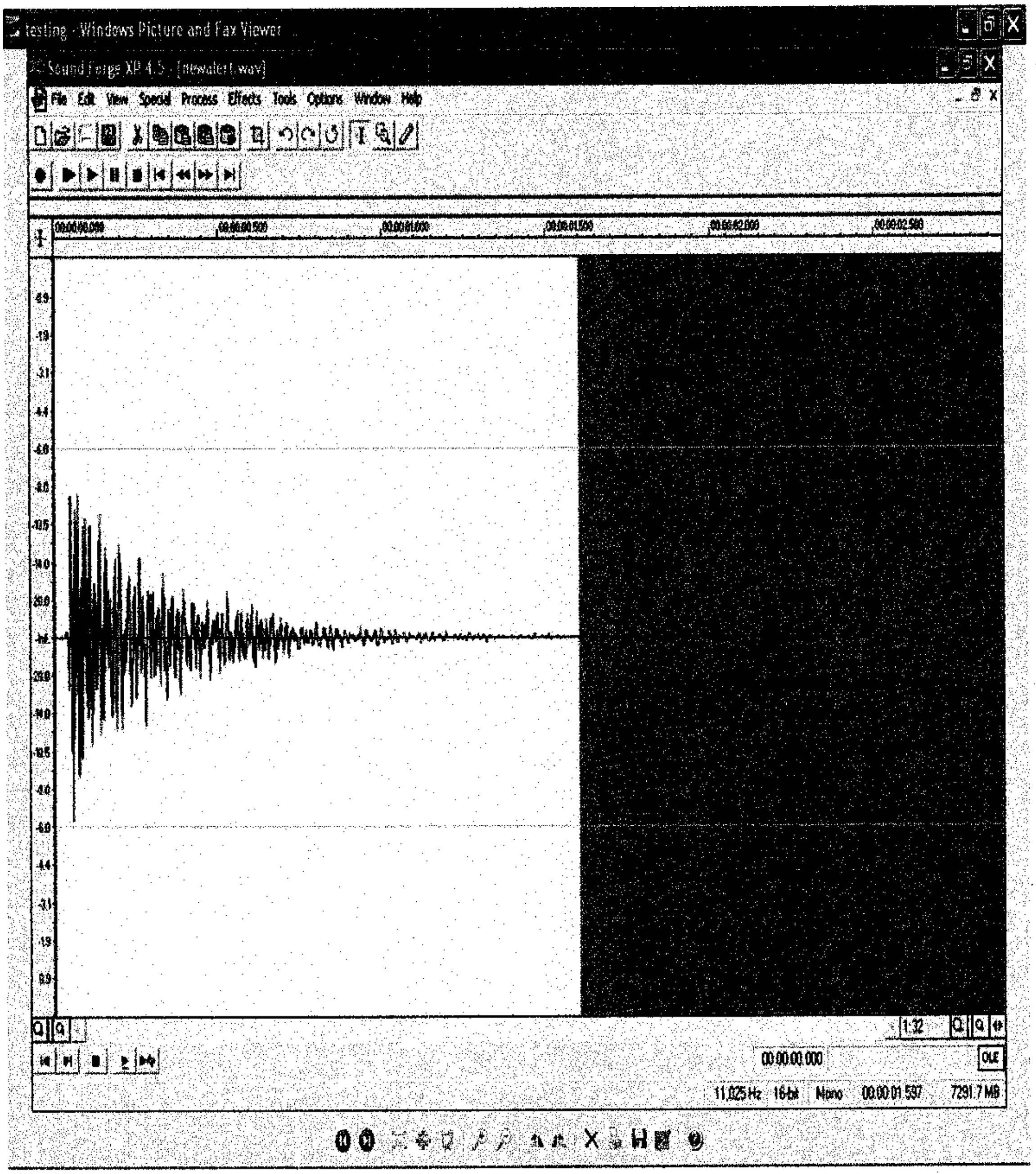


Figure 4.1 Example of wave sign from the original WAV file in the database

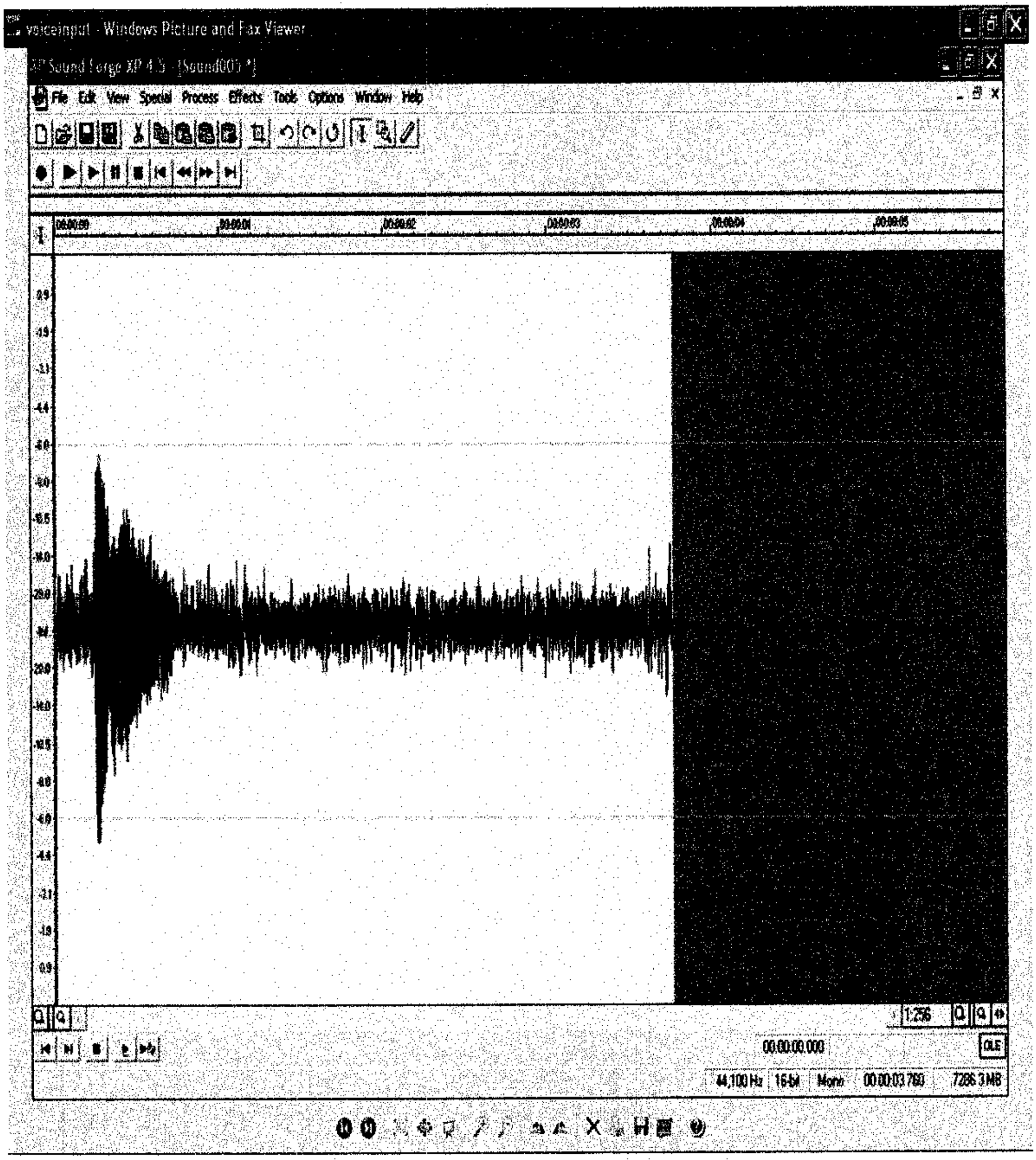


Figure 4.2 Example of wave sign from the input human voice

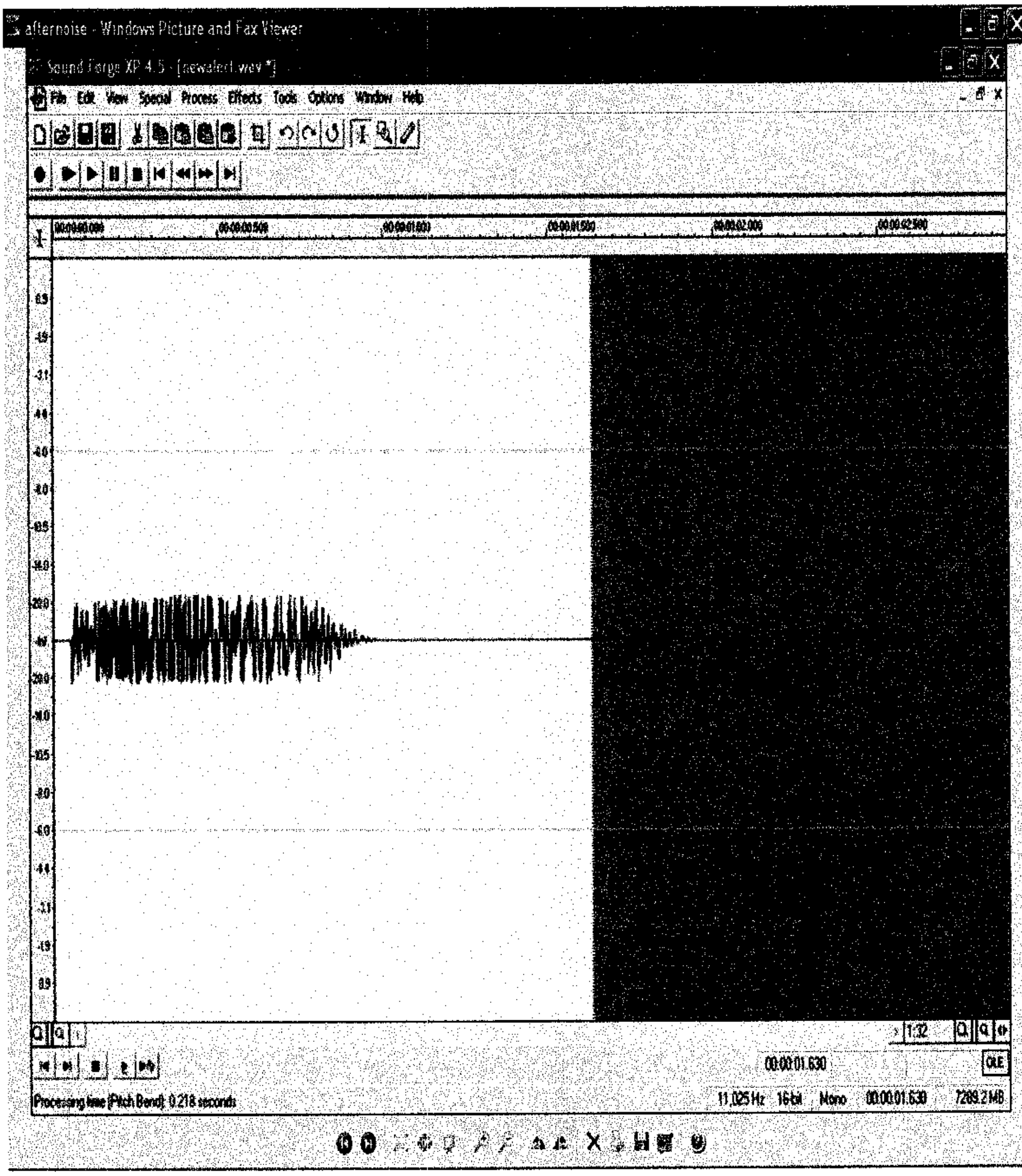


Figure 4.3 Example of wave sign after alteration are made

After proper research has been made, I now have found new software that is much more appropriate to my project that is Matlab. It provides the tools and function that make it possible for me to capture, compare and match the WAV file. I use the function of digital signal processing in Matlab to develop this project. Below is the example on how the process is done through Matlab.

First, I take same input and compared it with original song in database. Then each two files will be compared to output the error of comparison using the block parameter tool. The output of these two files is the plotting of wave sign showing the range of the error rate. The output will plot the graph based on the range of both files differences and also by both additional distraction after being compared. By looking at the graph, I then now write a function to find the matching title with the given output. The function is written based on the output graphs by setting the range of acceptable errors.

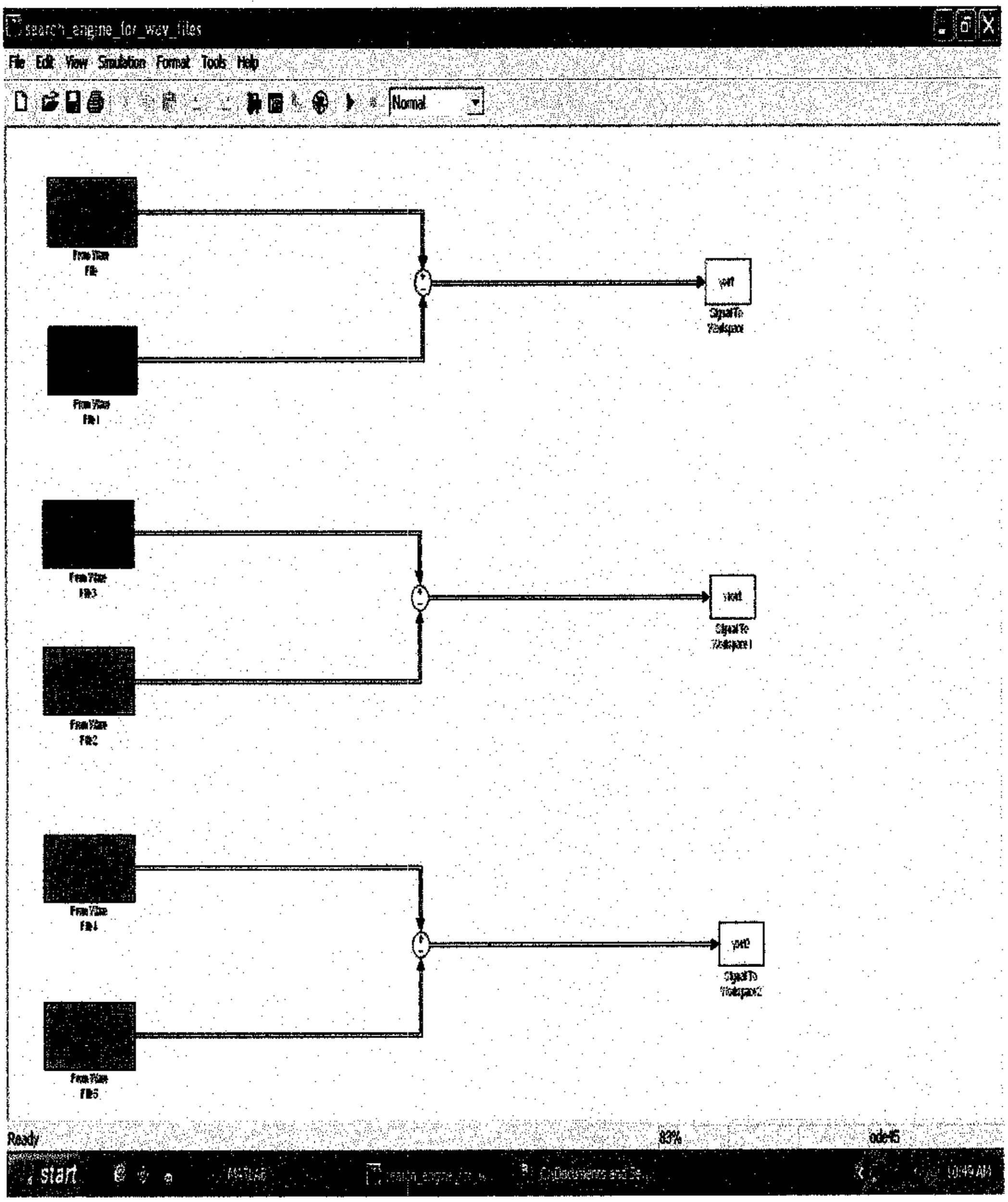


Figure 4.4 Simulink Model

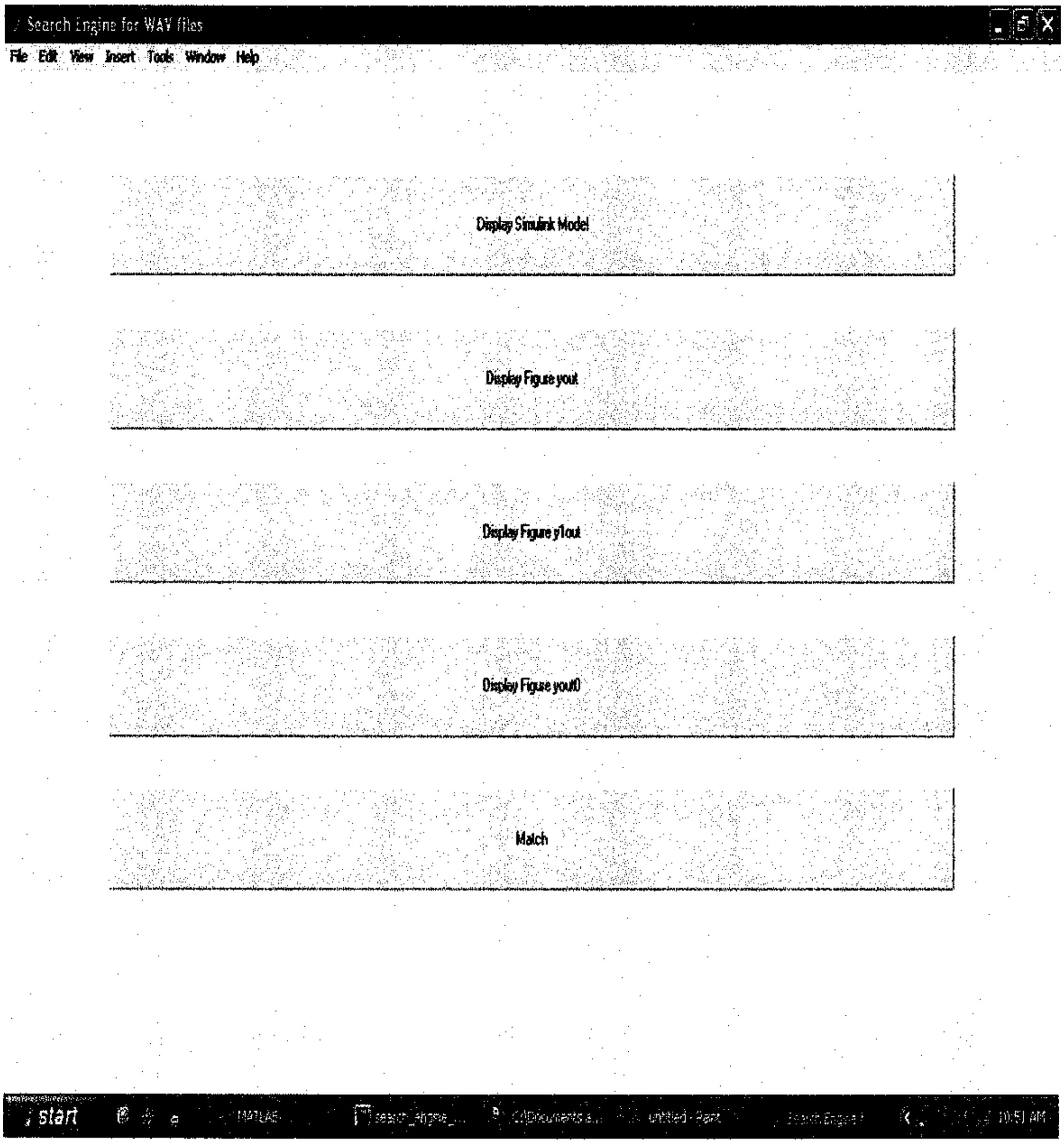


Figure 4.5 Graphical User Interface

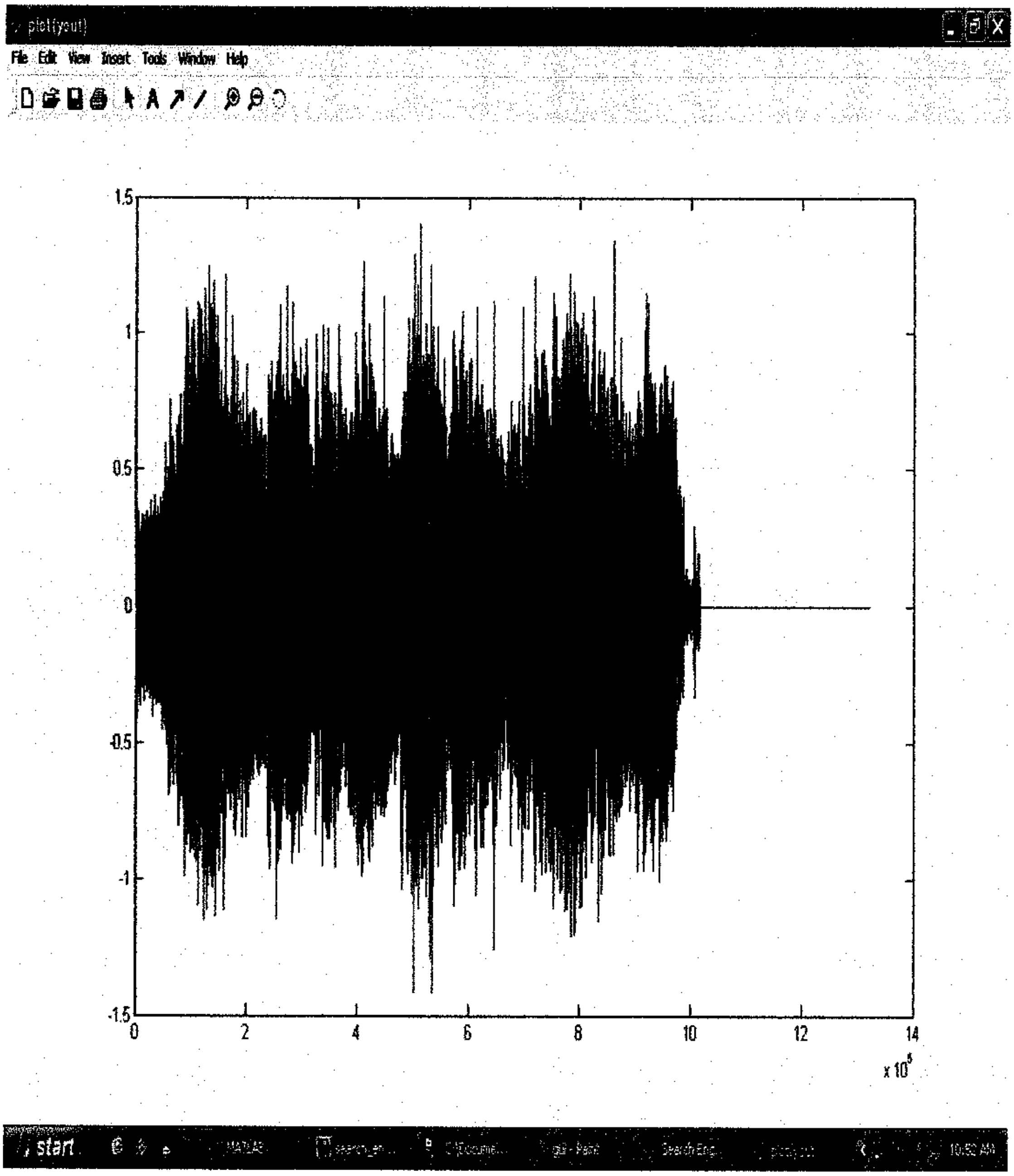


Figure 4.6 Error Graph Plotting (yout)

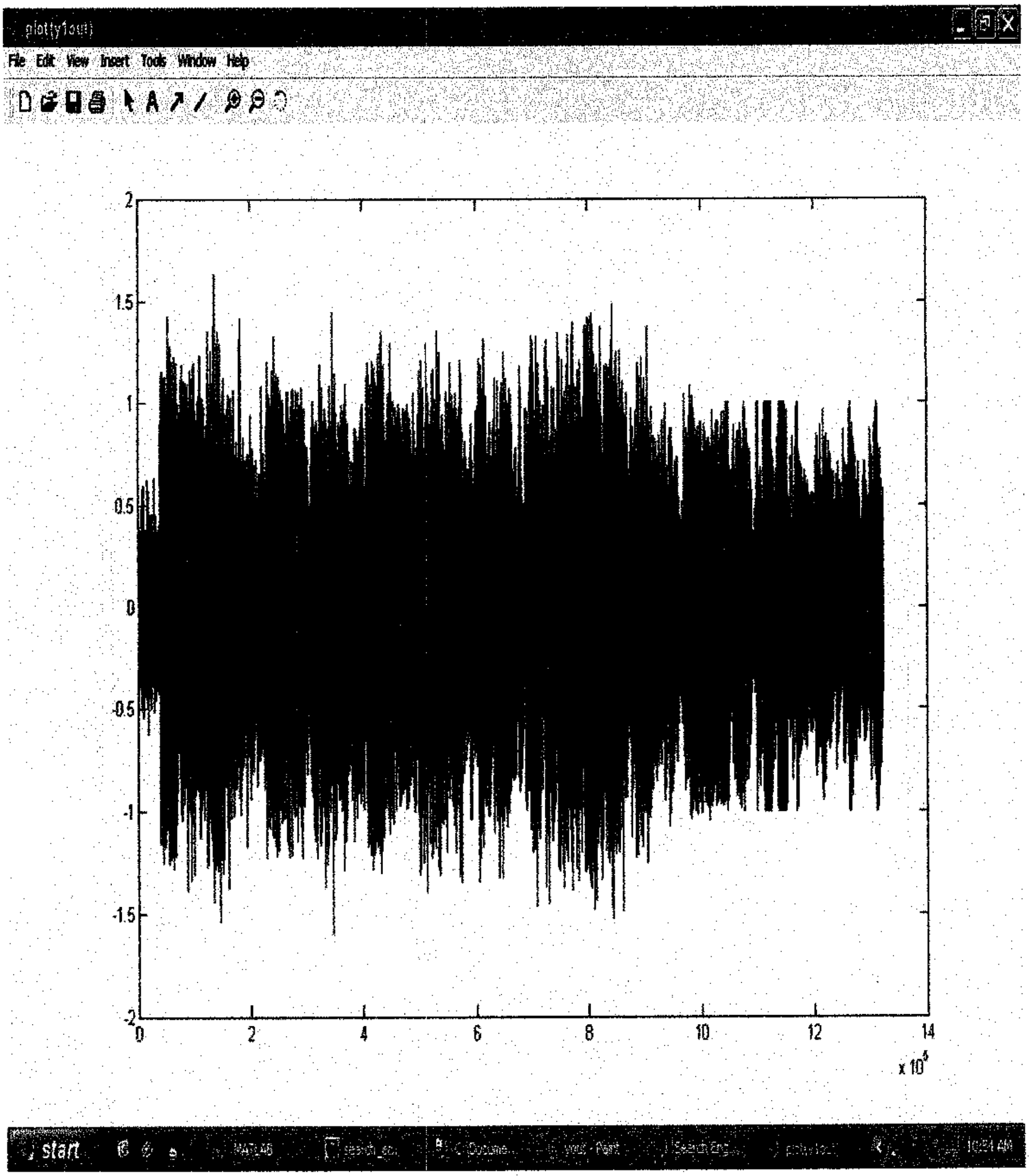


Figure 4.7 Error Graph Plotting (y1out)

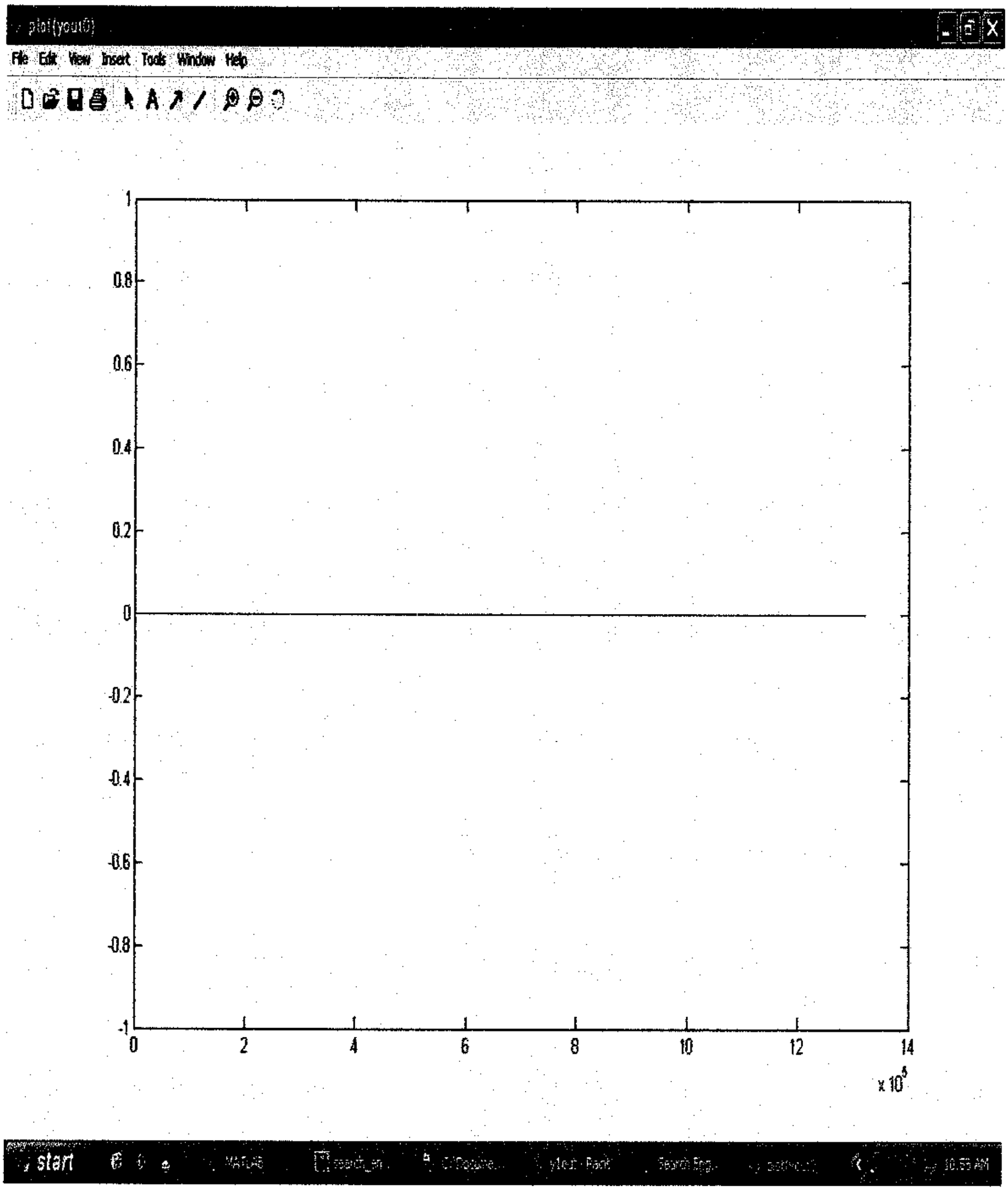


Figure 4.8 Error Graph Plotting (yout0)

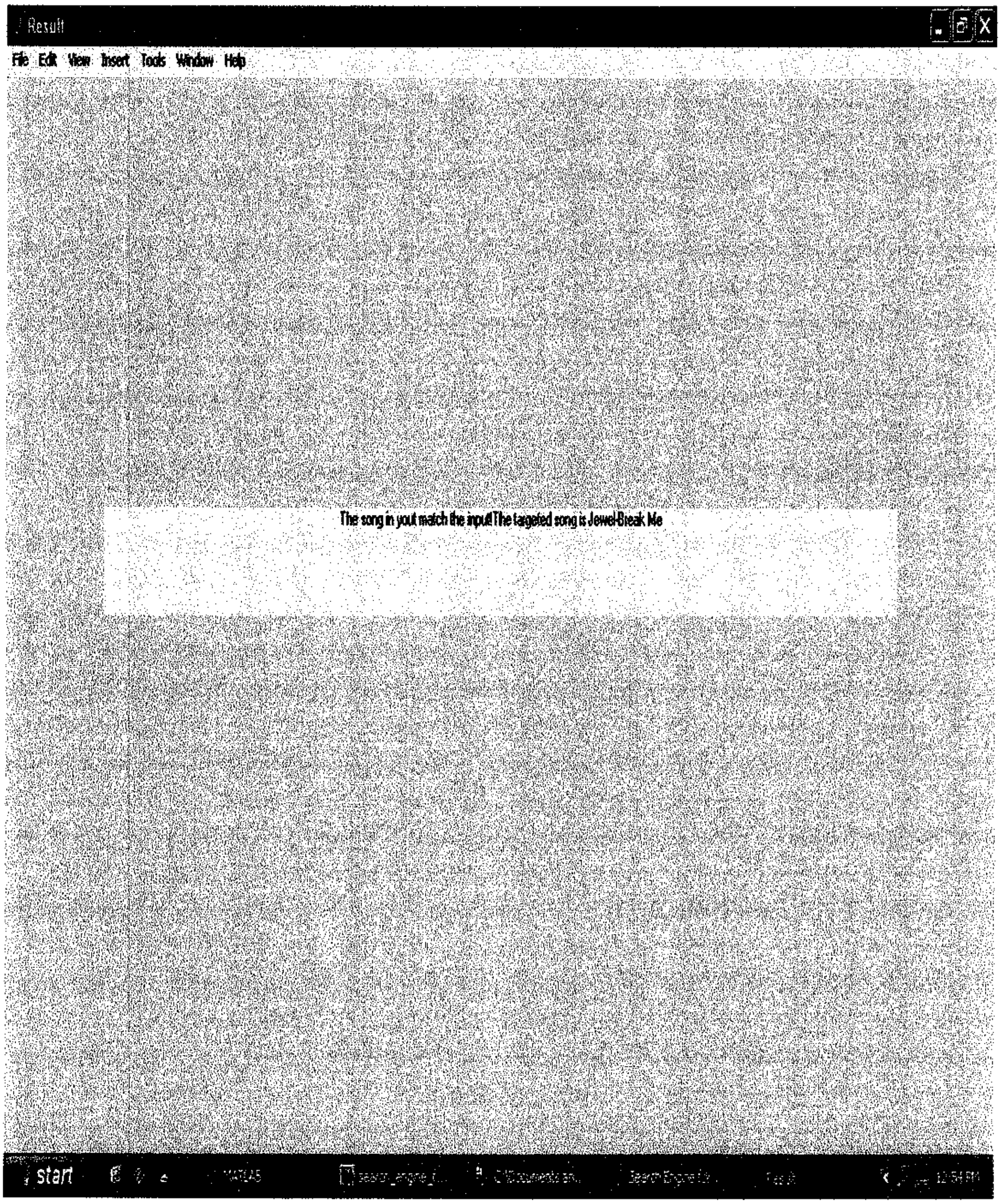


Figure 4.9 Result Display

Paperwork and also different studies being made by different people all around the world have provide me with the surface on how this project can be actually completed. Their research has yet to comply with my research as certain elements of it are some might and might not adaptable. The way they capture and convert the file might somehow differ to one another, but the concept similarity is still on the same track.

4.2 DISCUSSION

There are few consequences in executing my project. When user clicks the match button, the execution can take up to maximum of 7 minutes to display the results of matching title. It is because the computer I am using to execute this application is not a high-end computer but only a personal computer with moderate processor and RAM size. Executing MATLAB application involves simultaneous virtual memory processing tasks. Therefore, an appropriate allocation of virtual memory should be wisely allocated to run this application smoothly. This application also affected by the length of the portion song, the more length the longer it takes to display the result. This is because even in a 10 seconds portion of song, it may have up to 1,500,000 MB of frequency.

The research are mostly linger around the content-based music retrieval system, which the system are able to take the user's acoustic input (singing, humming, or music instrument playing, recorded original song played), but in my research it only concerns on the way those files can be compared with the original song in order to find the medium of similarity to match it. In the study also explained the pitch tracking methods and how the comparison engine works, and have given a mathematical analysis based on a probabilistic model to optimize parameters of the hierarchical filtering method. The response time and success rate of the systems demonstrate the feasibility of the project's usage as the query engine for music digital libraries.

A lot of research being made based on the online resources, journals, reference texts, and other related media such as newspapers and magazines have actually helped to strengthen my position in researching my project. The findings have resulted in an extraordinary research I ever made. Educating while applying what I found in a practical has helped me to be progressed in order to complete my project. Deeper understanding need to be done so that I will be able to take control the overall functions of WAV file format throughout completing the project and the practical work need also to be done in order to dig more into the project concepts and basis.

I have been advised to understand and study more on how WAV file actually can be manipulated. At the beginning of the progress I faced some obstacles especially in differentiating and manipulating between different WAV file, but after resource collection and finding being made, I am able to eliminate those problem. I also have to mangle around with different type of tools especially software tools that might be appropriate for the project development. For examples, SoundForge that provide basic tools and function to edit WAV files format. Furthermore, I have also learn digital signal processing majoring in audio by learning the functionality of Matlab.

Many progress can be seen and realize after proper study and research have been made based on the project. I can safely say that, I am on the right track in completing the project. I can safely say that, I have achieved all the objectives I set at the beginning of the project. The main goal is that, to be able to visualize and come out with the way to extract WAV file from database. This project is very essential for the understanding of WAV file manipulating and matching database. I learned a lot from this research project.

CHAPTER 5
CONCLUSION and RECOMMENDATION

Chapter 5 highlights the most significant findings in relation to the objective of the project. It actually summarize while conclude what I get from all the findings and the overall completion of the project development.

5.1 SUMMARY OF THE PROJECT

In this project, I had managed to achieve the initial objectives and successfully developed a search engine that has applied the task of capturing, comparing and matching of WAV files. I had successfully implemented the appropriate functions and tools for the user to ensure that they are able to use the application without any difficulties. Furthermore, from the research and study have been conducted throughout the project development, I found out that the project is capable to intensify the performance of task in extracting audio database by experimenting it using the Matlab functionality which utilizes the digital signal processing functions.

Moreover, the result from the research and study have also proved that performing sequence tasks in extracting audio database using the function of digital signal processing is very effective. Which mean that, I apply a step by step task such as modeling the simulink, inserting the block sum parameter then plotting the yout. It is shown that the implementation of sequence tasks in WAV file search engine is a successful performance altogether and it could be implemented in more future similar applications.

Throughout the accomplishment of this project, I had learn most of the functions and tools for developing a WAV file search engine while providing the ease of use to the user. From this, I had improved my ability on digital signal processing by theoretical and putting it into practice. In addition, I had also learnt on sequence tasks allocation in Matlab. Even though, I had faced many obstacles during development, I managed to complete the project on time. Also, I had learnt and improved the application of new and existing software throughout the completion of the project such as SoundForge and Matlab.

5.2 FUTURE RECOMENDATION

Although the project had been successfully completed, there are some propositions that could be implemented in the future for this prototype module. Due to time constraints and the fact of low and limited knowledge on subjects matter area, it is difficult to accomplish other important aspects that could be included in the project for a better productivity. Since the project that I am developing is would rather derive from a complex scope; I reckon that it is still the first step to develop it as a whole complex system.

What I am doing is the base for future development, as I will be providing the information and initial study on how comparison can be made between multiple types of WAV file inputs and the original intended song in the database and how this will lead to further matching. Besides that, I inspire audio developers to try out more digital signal processing functionalities as it is more attention-grabbing and utilize the computer audio advanced technologies.

In my application, not all steps are automatically executed, therefore I would like to suggest for future developer to try finding ways that can actually automate all the steps and make it more user interactive. I can't figure out on the ways to automated all the steps because it is due to come unavoidable constraints and not within the time and scope of study. As my application is just prototype that is build to help supplement and strengthen my research.

As mentioned earlier, this is only applicable in the computer database storage whereby, it only focuses on the WAV file format. Maybe in the future, it can then be applied on more advance platform such as the internet and can focuses on other audio file format such as mp3. It is also more effective if the application can be done using a better high-end technology and processor. I would also recommend that the comparing of the files can be made with the most accurate and error-free rather than putting consequences of error rate in the function. Although it might be complicated and complex to make the full system work as accurate as we want it to be, I would say that my project development is the stepping stone for other future work. In the future, students whom are interested to further continue this research paper should create more complex and challenging sequence tasks in various applications besides capturing, comparing, and matching the WAV files. Why do I say that it is complicated to accurately match? It is because, the complexity of the signal content such as the variation of pitch, frequency, and length that makes it difficult to 100% matched.

Therefore, I would conclude my finding and research as something that is not only educating people on how comparison and matching of WAV file can be done, but it also provide the crucial most needed resources and knowledge on how the project can be achieved successfully. As for that, I would give a credit for my hard work and full determination to complete this project.

5.3 CONCLUSION

As for the conclusion, the aim of this project is to provide user with the technology that will help them to search for WAV files that match the input given through the input device. Its objective is to be able to extract some audio database. It also had introduced the new method of extracting audio database and a new way to compare different WAV files. The findings of different materials have helped to boost the project realization and highlight its significant in relation with its objective. This project is to be a benchmark or guideline for developers to venture in depth in the area of digital signal processing and database extraction. The project had also contributed to a small portion of the digital signal processing and advanced technology world especially in audio field.

Although the development of audio search is relatively new, it is a direct information retrieval via audio mining that generally focuses on relatively noise-free, single-speaker recordings. I had applied the task of capturing, comparing, and matching of WAV files to achieve the objectives.

REFERENCES

1. **Ross Kirk and Andy Hunt,** "Digital Sound Processing for Music and Multimedia" 1999

2. University of Lincoln, Lincoln Media & Broadcast Prod. Centre

3. **Ghias, A. J. and Logan, D. Chamberlain, B. C. Smith**, "Query by humming-musical information retrieval in an Audio Database", ACM Multimedia '95 San Francisco

4. **McNab, R. J., smith, L. A. and Witten, Jan H**. "Signal Processing for Melody Transcription" Proceedings of the 19th Australian Computer Science Conference,1996.

5. **Jyh-Shing Roger Jang', Hong-Ru Lee** "Hierarchical Filtering Method for Content-based Musci Retrieval via Acoustic Input", Multimedia Information Retrieval Laboratory, Computer Science Dept, National Tsing Hua University, Taiwan

6. **John H.L. Hansen, J.R. Deller, and Jr., Michael S. Seadle.** "Transcript-Free Search of Audio Archives for the National Gallery of the Spoken Word", University of Colorado and Michigan State University

7. **John G. Ackenhusen,** "Real-Time Signal Processing", Prentice Hall PTR

8. **Glenn**, Ph. D, The Physics hypertextbook™,© 1998-2003 by Glenn

9. **BORG**, www.borg.com/~jglatt/tech/wave.htm

10. **Jakovljevic**<http://www.generalupdate.rau.ac.za/infosci/conf/Wed/Jakovljevic.htm

11. **Dena Taylor & Margaret Procter** Director & Coordinator Health Science Writing Centre, Writing Support and Coordinator, University of Toronto <http://www.doc.mmu.ac.uk/online/SAD/T02/sdlc.htm>

12. **Kathy Schwalbe**, Ph.D., PMP Augsburg College, Information technology Project Management 2nd Edition

13. **RusselKay**<http://www.computerworld.com.developmenttopics/development>May 14, 2002

14. **Prof. P.J. Ankiewicz & Dr. E. de Swardt** Department of Curriculum Studies Rand Afrikaans University <http://www.hyperdictionary.com/computing/SDLC>

15. **Voyetra Turtle Beach** <http://www.voyetra.com/site/default.asp>

16. **Bizness Online.com** <http://www.borg.com/~jglatt/tech/aboutiff.htm>

17. **Oreilly & Associates, Inc** <http://www.ora.com/centers/gff/formats/micriff/index.htm>

18. **Alan Dix, Janet Finlay, Gregory Abowd, Russel Beale.** Human Computer Interaction 2nd Edition Prentice Hall

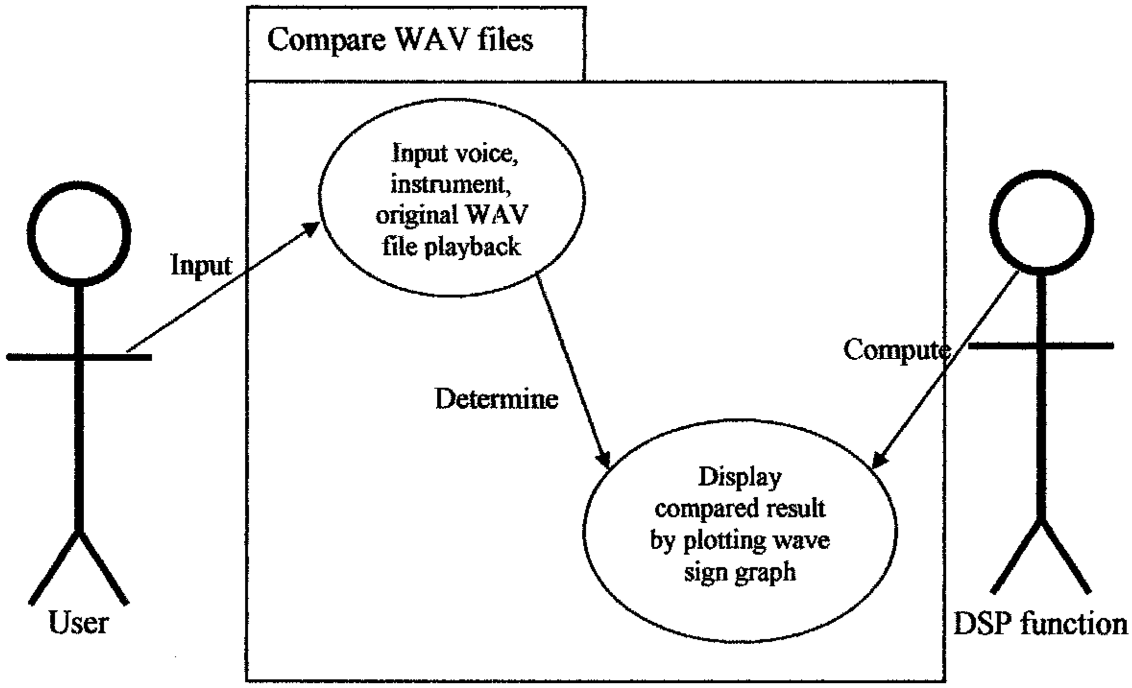
19. **Computer Speech Technology, Robert D. Rodman** Artech House Signal Processing Library

20. **Hilbert Transforms in Signal Processing, Stefan L. Hahn** Artech House Signal Processing Library

APPENDICES

Appendix 1

Use Case Diagram: Compare WAV files

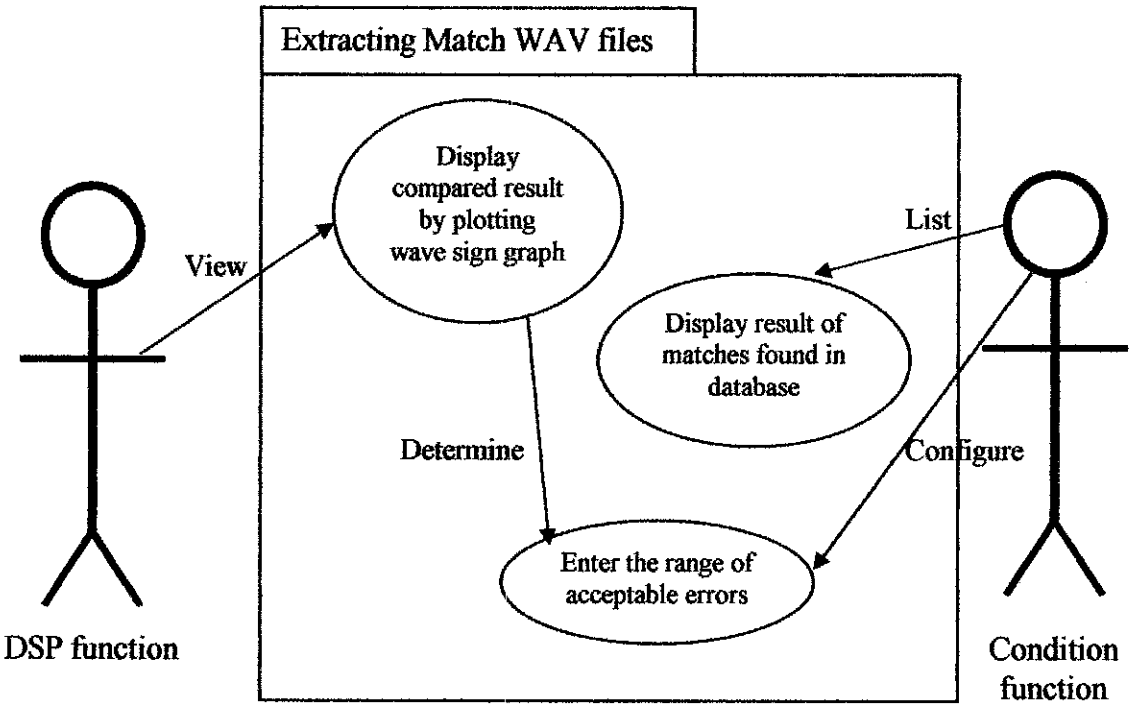


DSP = Digital Signal Processing

APPENDICES

Appendix 2

Use Case Diagram: Extracting Matched WAV files



Condition Function = MATLAB Programming Code

APPENDICES

Appendix 3

MATLAB Programming Code (M-File)

```
% Script : create a Graphical User Interface for Search Engine for WAV files
% In this mfile, a figure of 'a_win_f' is created.

%---------------------------------------------------%
% create 'a_win_f' Graphical User Interface   (main)   %
%---------------------------------------------------%

a_win_f=figure('Unit','normalized',...
  'Position',[0.35 0.30 0.30 0.4],...
  'Resize','on',...
  'Color','white',...
  'Numbertitle','off',...
  'Name','Search Engine for WAV files');

a_pb_sl=uicontrol('Style','Pushbutton',...
  'Unit','normalized',...
  'Position',[.1 .80 .8 .1],...
  'String','Display Simulink Model',...
  'Callback','search_engine_for_wav_files;');
%callback function: [search_engine_for_wav_files.mdl]

a_pb_fig=uicontrol('Style','Pushbutton',...
  'Units','normalized',...
  'Position',[.1 .65 .8 .1],...
  'String','Display Figure yout',...
  'CallBack','figure_yout;');
%callback function: [figure_yout.m]


a_pb_fig1=uicontrol('Style','Pushbutton',...
  'Unit','normalized',...
  'Position',[.1 .50 .8 .1],...
  'String','Display Figure y1out',...
```

```
  'CallBack','figure_y1out;');
%callback function: [figure_y1out.m]


a_pb_fig1=uicontrol('Style','Pushbutton',...
  'Unit','normalized',...
  'Position',[.1 .35 .8 .1],...
  'String','Display Figure yout0',...
  'CallBack','figure_yout0;');
%callback function: [figure_yout0.m]



a_pb_mth=uicontrol('Style','Pushbutton',...
  'Units','normalized',...
  'Position',[.1 .20 .8 .1],...
  'String','Match',...
  'Callback','compute');
%callback function: [compute.m]


%----------------------------------------------------%
%This program compute the range of acceptable error      %
%----------------------------------------------------%

a= length(yout)     %wrong length
b= length(y1out)    %wrong length
c= length(yout0)    %correct length


   SI1 = yout - yout0
   SI2 = y1out - yout0

   SI2 > SI1


%------------------------------------------------------------------%
%   create 'message_result' window to output the result of the match made   %
%------------------------------------------------------------------%


message_result=figure('Unit','normalized',...
     'Position',[0.35 0.30 0.30 0.4],...
     'Resize','on',...
     'Visible','off',...
     'Numbertitle','off',...
```

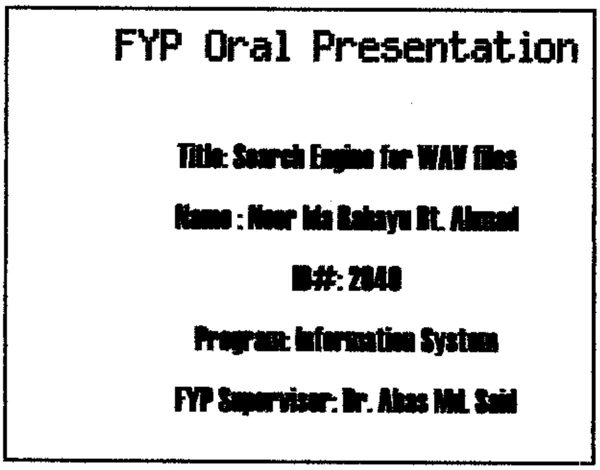
```
    'Name','Result');

   message_text_=uicontrol('Style','text',...
    'Unit','normalized',...
    'Position',[.1 .50 .8 .1],...
    'Visible','on',...
    'String','The song in yout match the input!The targeted song is Jewel-Break Me');


set(message_result, 'Visible' , 'on');




c_win_f=figure('Unit','normalized',...
  'Position',[0.35 0.30 0.30 0.4],...
  'Resize','on',...
  'Color','white',...
  'Visible','on',...
  'Numbertitle','off',...
  'Name','plot(y1out)');


plot(y1out,'color',[0 0 1]);




b_win_f=figure('Unit','normalized',...
  'Position',[0.35 0.30 0.30 0.4],...
  'Resize','on',...
  'Color','white',...
  'Visible','on',...
  'Numbertitle','off',...
  'Name','plot(yout)');


plot(yout,'color',[0 0 1]);


d_win_f=figure('Unit','normalized',...
  'Position',[0.35 0.30 0.30 0.4],...
  'Resize','on',...
  'Color','white',...
```

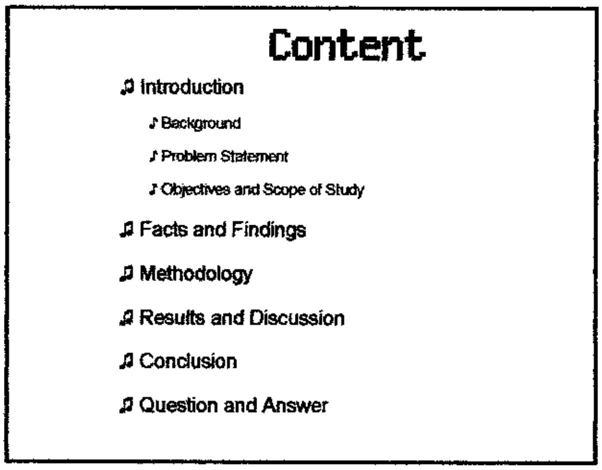
```
  'Visible','on',...
  'Numbertitle','off',...
  'Name','plot(yout0)');

plot(yout0,'color',[0 0 1]);
```

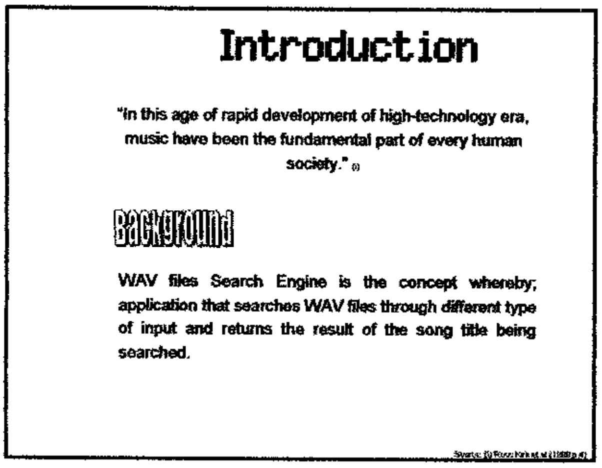
APPENDICES

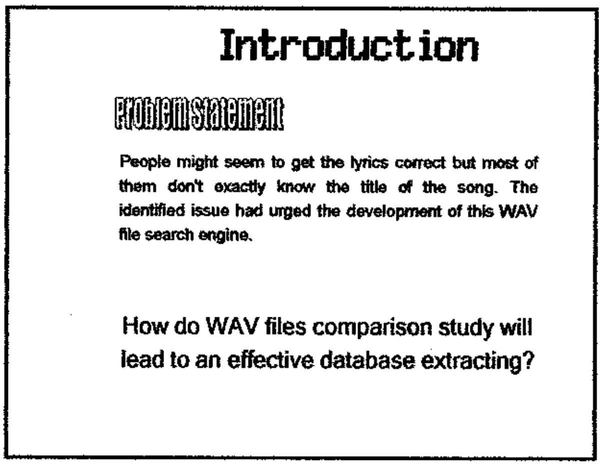
Appendix 4

Final Presentation Slide









Introduction

Objectives and scope of study

- To do research on the audio aspect of a WAV files
- To extract some of audio databases using specified method
- To show the different outcome of WAV files after comparison
- To understand the underlying concepts of search for WAV files

Facts and Findings

Theory outline

- Nature of sound
- WAV file format
- Audio Search Engine
- Digital Signal Processing
- Nature and content of signals

Facts and Findings

Nature of sound

"Sound is created by vibrating objects and propagated through a medium from one location to another." (ii)

- Sounds properties
  - Pitch and Frequency
  - Intensity and the Decibel Scale
  - Speed of sounds
  - Human Ear

WAV files format

Most common and probably the simplest sound formats, supported by Windows but consume large capacity.

Source: (ii) Lincoln Media and Broadcast Production Centre

Facts and Findings

Audio search engine

Mostly in the findings, the approach to the application is by formula and error consideration.

- Direct info retrieval via audio mining focuses on noise free, single speaker recordings. (iii)
- Specified algorithm for pitch tracking, based on energy levels and transition amount (length). (iv)
- Autocorrelation to obtain frequency, cut the pitch vector into notes. (v)

Source: (iii) John H.L. Hansen et al. (iv) R. J. McNab et al. (v) Chiao et al.

Facts and Findings

Digital Signal Processing

♫ Signal Processing is the process of extracting relevant information from the speech signal.

♫ Casting signal processing algorithms, both architectural design and algorithmic innovation are important.

Nature and Content of Signals

♫ Simpler signal is more preferable than the sound of spoken phrase; more understandable and also easy to develop in a simpler model.

Methodology

Project Framework

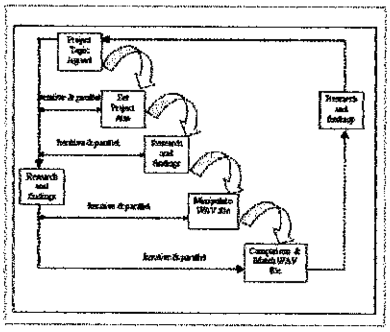
♪ No specific framework for describing the phases involved in developing this project.

♪ I don't practice a specific methodology which specifies which phase comes first and which comes next.

♪ Overall the project is actually done in parallel and without any indicated phases, but the concept of the project development flow is actually visible according to my progress.

Methodology

Project Workflow



Results and Discussion

Description

♪ take same input compare it with different database song

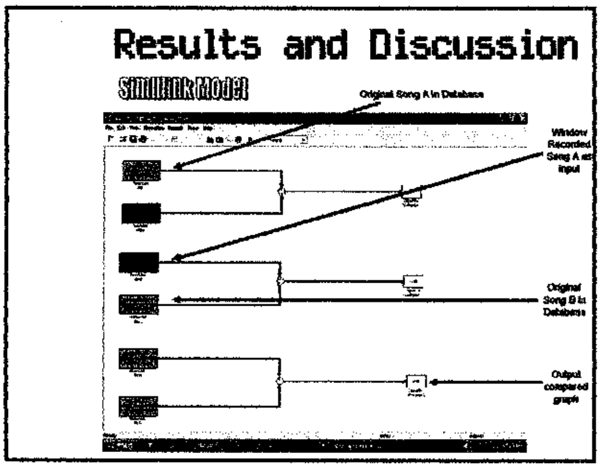
♪ uses the block sum parameter to calculate the error

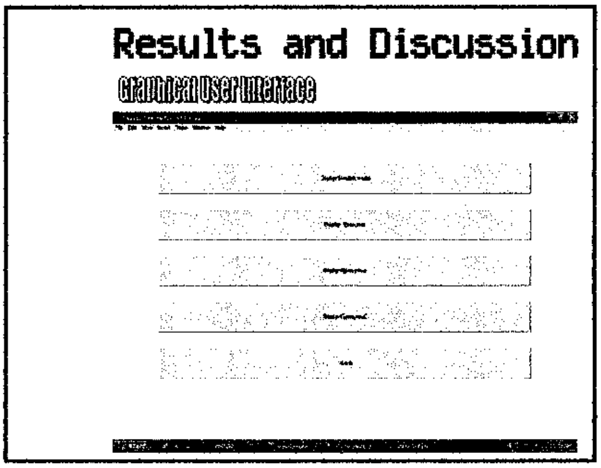
♪ plotting the compared error in graph form

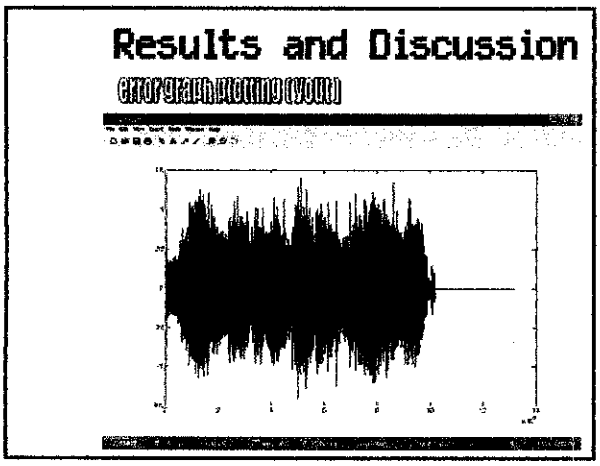
♪ match the given error graph by considering the range of acceptable errors

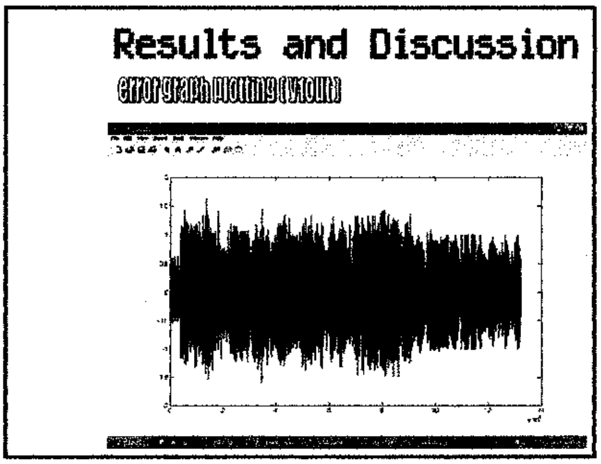
♪ output the result of match being made

♪ casting signal, both by architectural design and algorithmic innovation


Results and Discussion
Simulink Model
Original Song A in Database
Window Recorded Song A as input
Original Song B in Database
Output compared graph

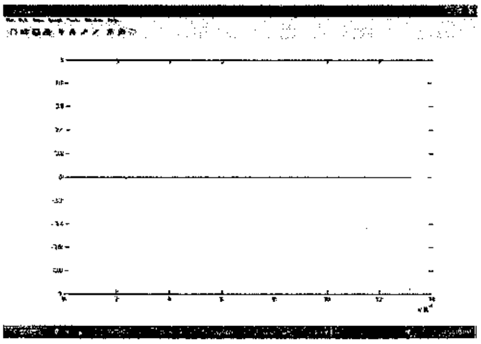

Results and Discussion
Graphical User Interface


Results and Discussion
error graph plotting (vout)


Results and Discussion
error graph plotting (v1out)

Results and Discussion

error graph plotting (violin)



Results and Discussion

Result Display

Results and Discussion

Discussions

- execution time can take up to 6 minutes to produce result
- computer need large RAM size because executing MATLAB involves simultaneous virtual memory processing tasks
- need to consider length of the portion song, the more length, the longer it takes to display result

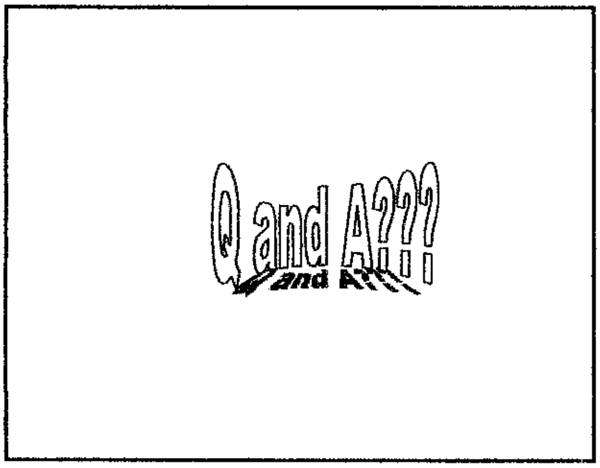
Conclusion

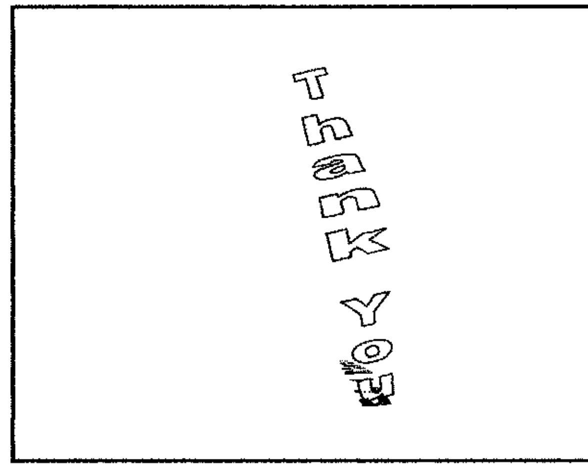
Conclusion

Although the development of audio search is relatively new, it is a direct information retrieval via audio mining that generally focuses on relatively noise-free, single-speaker recordings. I had applied the task of capturing, comparing, and matching of WAV files to achieve the objectives.

Future Recomendations

- applied in more advance platform such as internet
- focus on variety sound format such as .mp3
- applied in better end processor and RAM size


Q and A???


Thank You

APPENDICES

Appendix 5

Advertisement on Digi