WXES 3182

Zati Hakim Bt. Azizul Hasan
WEK 000492

System Login Authentication Using Voice Recognition And Other AI Methods

Supervisor : Mr. Woo Chaw Seng
Moderator : Mr. Nor Ridzuan Daud

ABSTRACT

Abstract

As computer becomes a commodity in work places and households appliances used by all – be it male or female, young or old –the issue of security and privacy are increasing from day to day. Today, every user wishes for a better way to protect sensitive data. For all the talk about beefing up system security, the most effective way to breach a user's personal files and applications is to guess his or her password. That's because in the interest of easy recall, familiar terms are most preferred by users in general when it come to remembering a password. In fact, most passwords can be found on paper scraps taped to the desk or under the keyboard.

This is where this project bores a significant impact on system security. Voice authentication, one of the most popular biometric methods in handling fraud and data intruders, perhaps is the very thing that could change everything... Still one of the leading Exciting areas of modern computer science research, voice recognition and authentication is also one of the most difficult. The sheer variety and complexity of a word makes recognizing similar words very difficult even though it is said by the same individual. Nevertheless it is still the cheapest and yet effective approach to user's authentications because of the simplicity of the devices used – everyone has a microphone or a telephone attached to their own computer and everyone has a unique vocal tract that cannot be changed.. Whenever there is need for security, there exist an opportunity for this unique technology It's up to us to accept its uniqueness and ubiquity - a powerful combination that can further allow for powerful solution.

ACKNOWLEDGEMENT

Acknowledgement

I am most grateful to Allah SWT, for providing me the strength to keep carry on, to forget giving up, and to always believe success is sweetest when it is well-earned.

It is not possible for me to name all of the many who supported the development of this project, however, certain individuals made special contributions and they deserve special recognition.

First and foremost, my utmost gratitude to my supervisor, Mr. Woo Chaw Seng; whose innovative ideas and intelligent views have helped me defined a more specific project scope and objectives. Aside from that, his constant guidance and understanding throughout the project duration has enabled me to clear any insecurities and misunderstandings about the project development. This project would have not even taken place if my supervisor had not initiated it. Thank You Mr. Woo.

A big thank you too to my thesis moderator, Mr. Md. Nor Ridzuan Daud, whom gave insightful views about enhancing the project during the VIVA presentation. Some ideas have been considered in the system design phase and the rest were also considered during the implementation phase.

My scientific colleagues, Shamsul Adli B. Adlan and Metrius Benedict; have without exception been most helpful. Special thanks to Mohd. Shahman B. Samsul Ambia, once only an acquaintance, now a close friend. Without their co-operation,

unending supply of bright ideas, and most of all understanding, this project would have been a disappointment. You guys are the best!!!

A tribute too to Suryani Ahmad; I may have lost you forever dearest friend, but somewhere deep inside, your voice still lingers, an inspiration – I will always hold onto the memories we shared as long as I live. Miss you …

Last but never least my family, especially my parents, whom gave me boundless support and firm believe throughout these two semesters. My deepest appreciation, Mak, Abah, everybody, for everything you are to me. Alhamdullilah with all the support I was showered, this documentation is accomplished and I end this with the most heartfelt THANK YOU.

TABLE OF CONTENT

Table Of Content

Chapter 7 – Evaluation

Chapter 8 – Conclusion

Appendices

LIST OF FIGURES

AND TABLES

List of Figures :

List Of Table

CHAPTER 1

-INTRODUCTION-

1.1 Project Background

In order to complete the Bachelor's Degree in Computer Science, undergraduates are required to accomplish the final year project, which is to build a working system that encompasses skills and knowledge learnt in the past semesters. Therefore, three fellow undergraduates and I are building a login system that uses voice recognition to authenticate users entries using various AI methods. As mentioned the title of this project is ***'Login Authentication Using Voice Recognition And Other AI Methods'***. The main objective of this project is to create and develop a more innovative and secure login system to access a personal even a networked computer system compared to current ones available for Malaysians. By applying various AI techniques like Artificial Neural Network (ANN), for example, the system will be able to provide a new approach to recognize and validate an individual through his or her voice. Essentially, it is hoped that with the aid of this system the users may gain a better security comfort knowing their privacy and priority are better protected.

1.2 Project Definition

1.2.1 Definition of Login System

A login system means the account name used to access a networked computer system. It is the first thing any user would have to pass in order to access the computer system. Usually it requires an ID or an account name plus a specific password that only belongs to the user. These two are then keyed in using keyboard and access is granted for users

with the correct match of IDs and passwords. [J.Radcliffe (1997), Webster's Pocket Computer Dictionary, DJF International]

1.2.2 Definition of Authentication

The ability to validate that an individual is actually the person with whom a system is communicating or conducting a transaction is called authentication. Authentication is accomplished using one or more of three validation approaches: knowledge factor (something the individual knows), possession factor (something the individual has), or a biometric factor (something physiologically unique about the individual).

[J.Radcliffe (1997), Webster's Pocket Computer Dictionary, DJF International]

1.2.3 Definition of Voice Recognition

Voice recognition is the technology by which sounds, words or phrases spoken by human are converted into electrical signals, and these signals are transformed into coding patterns to which meaning has been assigned. The most common approaches to voice recognition can be divided into two classes: identification (feature analysis) and verification (template matching). [J.Radcliffe (1997), Webster's Pocket Computer Dictionary, DJF International]

1.2.4 Definition of System Login Authentication Using Voice Recognition

System Login Authentication Using Voice Recognition is a system that has the ability to access a networked computer system, using an ID or an account name plus a specific biometric (voice) password that only belongs to the user. It also has the ability to authenticate that an individual is actually the person with whom a system is communicating with by verifying his or her voice***1.3 Project Overview***

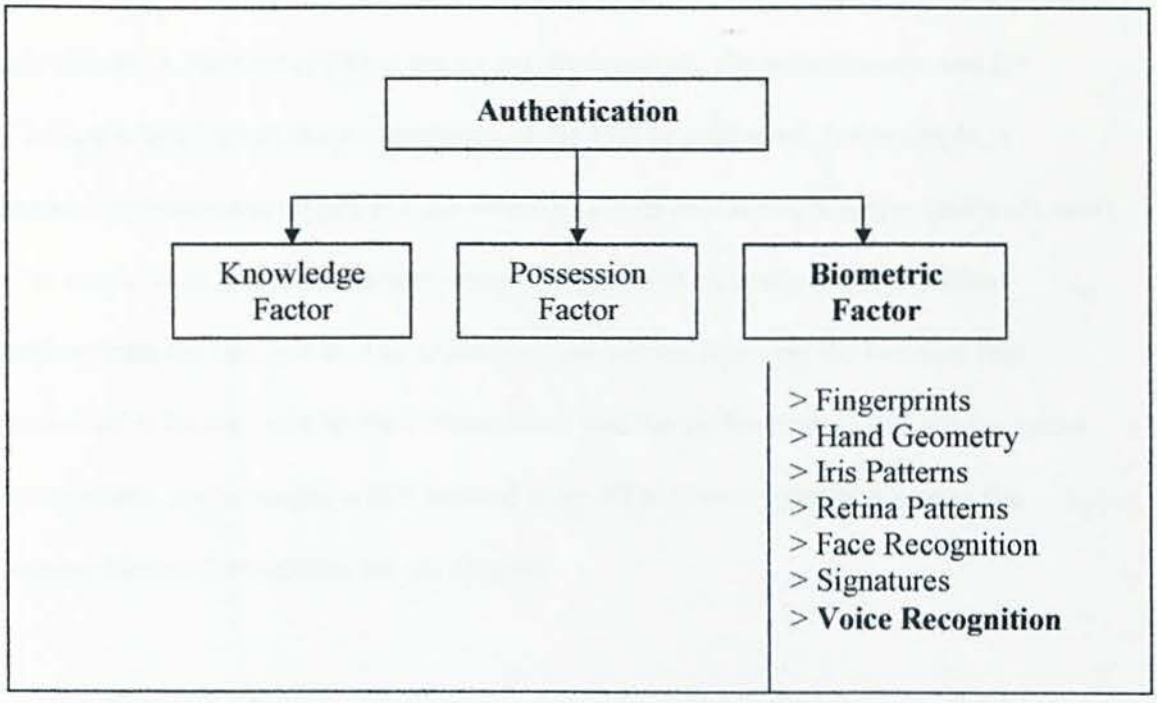


Figure 1.1 Three Approaches To System Authentication

For this paper, I will be referring to biometric the most, focusing specifically and fully on the technology of voice recognition.

1.3.1 Overview of Authentication

As mentioned in section 1.2.1, authentication can be done in three different ways : using knowledge factors, personal factors and also biometric factors. Next are brief descriptions for the first two.

Knowledge factors are something an individual "knows," such as a Personnel Identification Number (PIN) or password. Both parties, the authenticator and the "authenticatee," must share knowledge of the PIN or password. For example, a person (authenticatee) types in a password to log on to a network server (authenticator). The server must know the person's password in order to verify it and therefore authenticate the individual. The security of the system relies on the fact that that password is known only by the authenticator and the authenticatee, and is kept secret from others. For example, a PIN entered at an ATM is encrypted and sent to the issuing financial institution for verification.

Possession factors are something an individual "has," such as a door key, an employee badge, and a cryptographic key. When symmetric keys are used for authentication, typically the authenticatee creates a cryptically derived check value, called a Message Authentication Code (MAC), that the authenticator can verify using the same key. Thus, the authenticatee and the authenticator must share the symmetric key, but neither party actually knows what its unique property or identifier is. The security of the system relies on the fact that, unlike knowledge factors, the specific contents of a

symmetric key are kept absolutely secret, even from the users of the key. This makes the initial generation and exchange of the symmetric key a bit tricky, which requires key management techniques that are beyond the scope of this particular paper. Asymmetric keys, often referred to as public-key cryptography, or Public-Key Infrastructure (PKI) technology, are another form of cryptography used as a possession factor. In this scenario, the authenticatee has possession of an asymmetric private key and the authenticator has possession of the corresponding asymmetric public key. The authenticatee creates a digital signature using the private key against a same key, and the security of the system relies on the fact that the private key is kept absolutely secret. Only the authenticatee has access to using the private key. In addition, the integrity and authenticity of the public key must be established and maintained using a technique called digital certificates, which requires certificate management techniques that are also beyond the scope of this paper.

The third approach to authentication is described next under a new subtitle because of its significant importance to the project.

1.3.2 Overview of Biometric

A biometric factor is something physiologically unique about an individual, such as a fingerprint, facial image, iris scan, voice pattern, and handwriting (as stated in Figure 1.3). Many other types of biometric technology have been developed. When an individual wants system access, a sample is taken of the authenticatee's biometric data, for example, a digitized signature. Then, the authenticator, using a previously enrolled

version of the same biometric (called a template), can match the sample against the stored template to verify the individual's identity. Biometrics are not secret, as everyone leaves finger-prints everywhere they go, faces and eyes can be photographed, voices can be recorded, and handwriting samples can be obtained. The security of the system therefore relies on the integrity and authenticity of the biometric information, which can be accomplished once the individual has been enrolled.

I will elaborate more on this fascinating field of biometrics in The Literature Review (Chapter Two).

1.3.3 Overview of Voice Recognition

Voice recognition is a computer application that lets people control a computer by speaking to it. In other words, rather than using a keyboard to communicate with the computer, the user speaks commands into a microphone (usually on a headset) that is connected to a computer.

By speaking into the microphone, users can do two things. First, they can tell their computers to execute commands such as open a document, save changes, delete a paragraph, even move the cursor--all without touching a key. Second, users can write using voice recognition in conjunction with a standard word processing program. When users speak into the microphone their words can appear on a computer screen in a word processing format, ready for revision and editing.

1.4 Project Objectives

Objectives are defined as goals or aims. Hence these project objectives listed below are identified as the aims hoped to achieve by developing this system.

- **to create and develop a more innovative and secure login system to access a personal even a networked computer system .**
- **to build a system that can authenticate a user by verifying his or her voice regardless of gender and age.**
- **to explore and apply AI techniques in voice recognition.**
- **to expand feature of a computer system with minimal usage of keyboard.**

1.5 Project Scope

In this section of the documentation, the project scope would be defined. The project scope explains what the system will do and what it will not.

- **System will feature a very simple and straightforward user interface**
- **There will be no nice user interface**
- **System results will be a binary output – meaning system will only generate a 'Yes' or 'No' answer to show user's verification/authentication, rather than outputting percentage or probability of the match.**

- **Only recorded voice is acceptable as system's input. System will not tolerate other sources, example music or long speeches.**
- **All users must have an ID registered before voice enrollment .**
- **All first-time users must enrolled at least once.**
- **Regular users does not required to register or enroll their voice more than once unless they feel like changing to a new password.**
- **System can do a simple microphone testing before each voice recording to reduce recording error.**
- **To train the network, system must accept few enrollments from each user.**
- **Playback is available if user wants to confirm their enrollments.**
- **To verify, system will do voice recognition by matching user's live template with their respective voice templates stored previously in the network.**
- **System will only verify a user not identify a user.**
- **System will impost the human-friendly image because the authentication of users done via voice – a feature related to human natural behaviors.**

1.6 Project Schedule

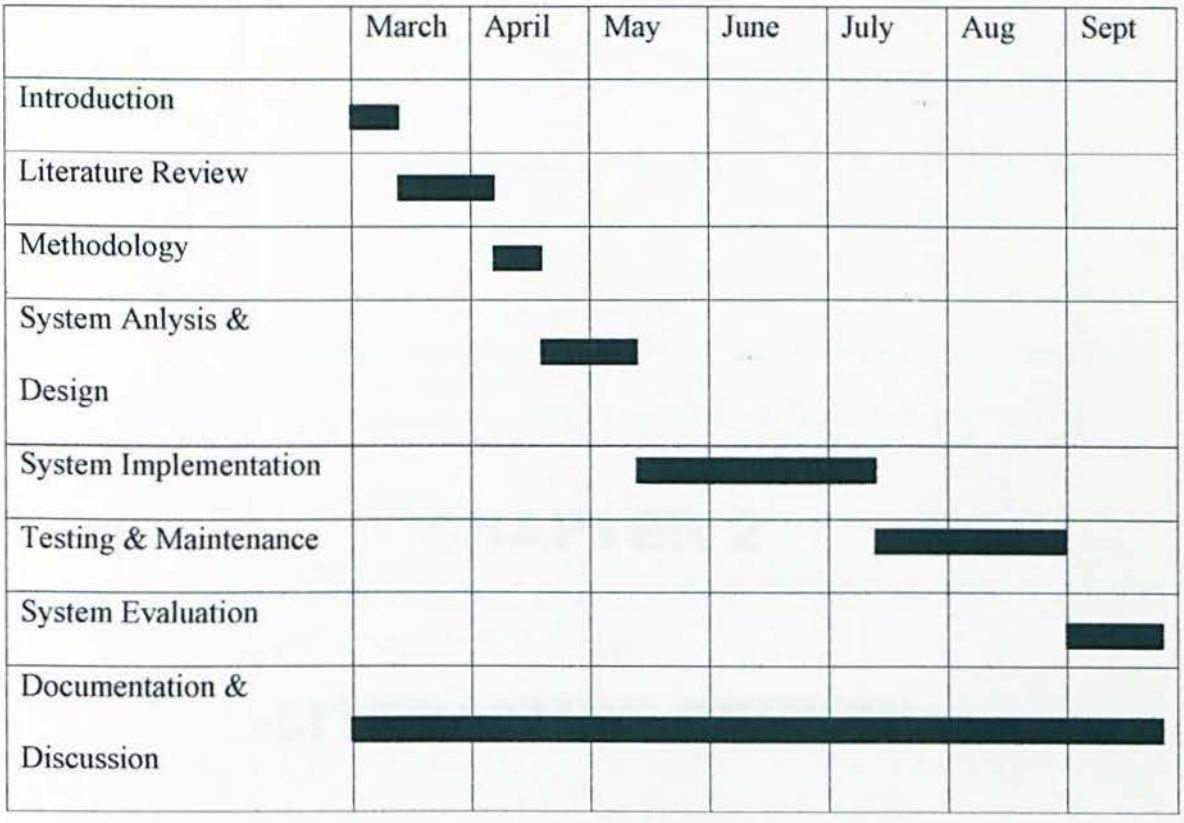


Figure 1.2 Project Schedule

1.7 Chapter 1 Summary

This chapter introduces the project background , an overview of what the system is all about in general. Among others it touches the concept of authentication, biometric and voice recognition in order to give the viewers a general light on the subjects. Chapter Two, next , is a continuation of study of these subjects in detail, plus further analysis on several knowledge domains.

CHAPTER 2

-LITERATURE REVIEW-

2.1 Overview of Literature Review

A literature review is an account of what has been published on a topic by accredited scholars and researchers. It is more often included in a research report, essay or thesis. The purpose is to express knowledge and ideas established on a research or thesis topic, and to convey their strengths and weaknesses.

Writing a literature review would allow me to gain and demonstrate skills in two areas:

1. **Information seeking**: the ability to scan the literature efficiently, using manual or computerized methods, to identify a set of useful articles and books
2. **Critical appraisal**: the ability to apply principles of analysis to identify unbiased and valid studies.

In this section I hope to fulfill these criteria listed in accomplishing a good review:

- Be organized around and related directly to the thesis or research question being developed.
- Synthesize results into a summary of what is and is not known
- Identify areas of controversy in the literature
- Formulate questions that need further research

2.2 The Study Of Biometric

Biometric : How It Work

The single data representation of a biometric characteristic or measurement derived from an individual's fingerprint, voice, iris, face, or handwriting, which is captured or scanned by a biometric device, is called a *biometric sample*. The information extracted from one or more biometric samples is used to create a *biometric template*. An individual is authenticated when a current *biometric sample* is found equivalent to, or "matches," the *biometric template*. Both the *biometric sample* and the *biometric template* are called *biometric data,* or *biometric information*. An automated system capable of collecting, distributing, storing and processing biometric data, and returning a decision (match or non-match), is called a *biometric system*. A typical authentication process utilizing biometric technology consists of the following basic steps:

1. Capture the biometric data using a physical reader device;
2. Evaluate the quality of the captured biometric data and recapture if necessary;
3. Process the captured biometric data to create a biometric sample;
4. Match the biometric sample with a previously enrolled template, or templates, to determine if a match exists. This matching can be done as verification or identification.

These steps utilize three fundamental biometric applications: Enrollment, Verification, and Identification.

Enrollment

Enrollment is the process of entering a new biometric template and identifier into the database. This data is usually entered along with other information about the individual, which links the individual to an organization, an account, or a set of privileges. Enrollment can incorporate *identification* to make sure that the individual is not already in the database, perhaps under another name. If no match is found, the biometric template, the identifier and its associated information can be added to the database.

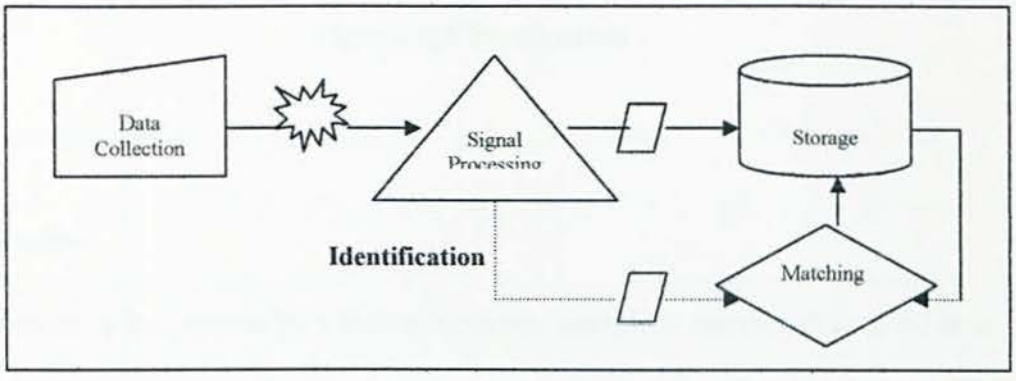


Figure 2.1 Enrollment

Verification

Verification involves a "one-to-one" comparison of a current biometric sample with a particular, previously generated biometric template, stored in a database or on an ID card, in order to ensure the correctness of the user's "claimed identity." The biometric template is retrieved from the database using the user's claimed identity, indicated by the user ID, user name, etc., or it is assumed based on the user's possession of the ID

card containing the biometric template. If the biometric sample matches the previously generated biometric template, then the claim of identity is verified.

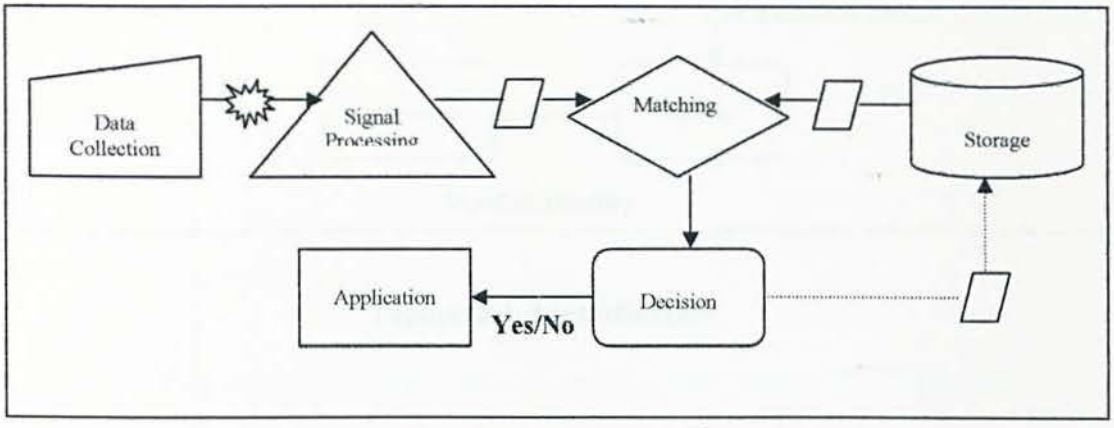


Figure 2.2 Verification

Identification

Identification is the process by which a biometric sample is compared with *all or a specified subset* of the enrolled biometric templates, or a subset based on search algorithms, in the database. This "one-to-many" comparison is done in order to find a matching template and thus identify the person who provided the biometric sample. Unlike verification, the user does not provide a "claimed identity" but instead is identified strictly on the basis of the biometric sample matching one of the biometric templates in the database. The technique can be used for recognition or to confirm that the person being identified is not known under a different name.

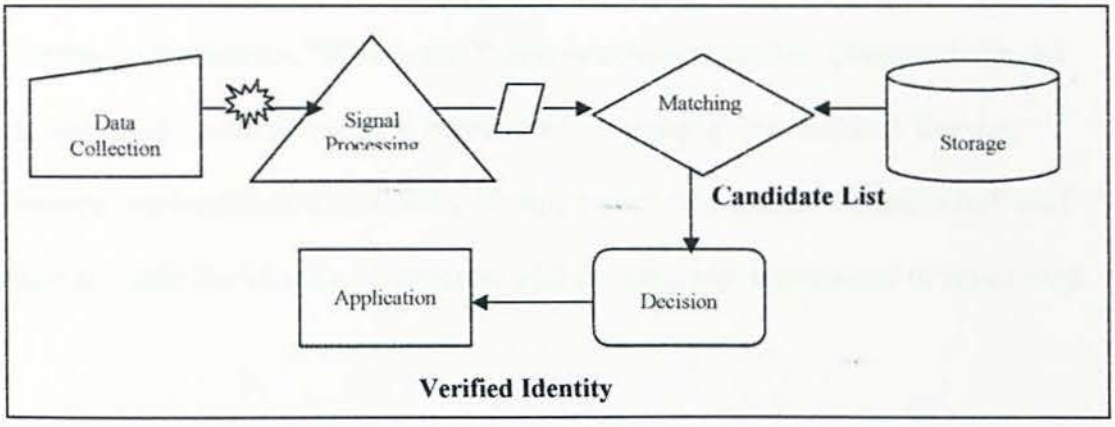


Figure 2.3 Identification

Biometrics : Identification Vs Verification

In the biometrics industry, a distinction is made among the terms identification, recognition and verification. In both identification and recognition, essentially synonymous terms, a sample is presented to the biometric system during enrollment. The system then attempts to find out who the sample belongs to, by comparing the sample with a database of samples in the hope of finding a match (this is known as a one-to-many comparison). Verification is a one-to-one comparison in which the biometric system attempts to verify an individual's identity. In this case, a new biometric sample is captured and compared with the previously stored template. If the two samples match, the biometric system confirms that the applicant is who he/she claims to be. The same four-stage process - capture, extraction, comparison, and match/non-match - applies equally to identification, recognition and verification. Identification and recognition involve matching a sample against a database of many, whereas verification involves matching a sample against a database of one. The key distinction between these two approaches centers on the questions asked by the

biometric system and how these fit within a given application. During identification, the biometric system asks, "Who is this?" and establishes whether a biometric record exists, and, if so, the identity of the enrollee whose sample was matched. During verification, the biometric system asks, "Is this person who he/she claims to be?" and attempts to verify the identity of someone who is using, say, a password or smart card.

Biometric Requirements

By its very nature, biometric information cannot be kept secret. Individuals leave finger-prints, show their faces, use their voices and leave samples of handwriting behind. Thus, an authentication system should not rely on the confidentiality of biometric information because confidentiality of biometric information cannot be achieved. However, where biometrics are linked to individual identity, privacy requirements may require that such data is encrypted to prevent its disclosure to those who are not authorized. An authentication system based on biometrics relies on the integrity and authenticity of that biometric information. Thus, biometric information must be protected against unauthorized modification and substitution. Furthermore, the source and destination of the biometric information must be protected to prevent biometric data from being "injected" into the system. Such protection should include temporal information to eliminate falsified or replayed data. Accordingly, the primary requirements are to establish and maintain the integrity and authenticity of biometric information during transmission and in storage. Authenticity and integrity of biometric information can be achieved using physical protection where no transmission is involved and all biometric components reside within the same tamper resistant unit. When transmission occurs, cryptographic mechanisms are the only

viable alternative, since physical protection is typically not practical. Other considerations for managing and securing biometrics systems include metrics such as the failure to enroll rate, false match rate, false non-match rate, and various types of threats such as identity theft, hill-climbing attack and synthetic attack, all of which are beyond the scope of this paper.

Biometric : Authentication Methods

Custom to those briefly mentioned in Chapter One, I will present here a more detailed description on each and every authentication methods of biometric except for voice recognition, as it bores a significant importance to this project and will be attended to under another subtitle.

- Fingerprint

  In recent years, fingerprints have rallied significant support as the biometric technology that will probably be most widely used in the future. In addition to general security and access control applications, fingerprint verifiers are installed at military facilities, including the Pentagon and government labs. Although machines tend to reject over 3% of authorized users, the false acceptance rate (FAR) is less than one in a million. Today, the largest application of fingerprint technology is in automated fingerprint identification systems (AFIS) used by police forces throughout the U.S. and in over 30 foreign countries. The fingerprint's strength is its acceptance, convenience and reliability. It takes little time and effort for somebody using a fingerprint identification device to have his or her fingerprint scanned. Studies have also

found that using fingerprints as an identification source is the least intrusive of all biometric techniques.

Verification of fingerprints is also fast and reliable. Users experience fewer errors in matching when they use fingerprints versus many other biometric methods. In addition, a fingerprint identification device can require very little space on a desktop or in a machine. Several companies have produced capture units smaller than a deck of cards. One of the biggest fears of fingerprint technology is the theft of fingerprints. Skeptics point out that latent or residual prints left on the glass of a fingerprint scanner may be copied. However, a good fingerprint identification device only detects live fingers and will not acknowledge fingerprint copies.

- Hand Geometry

  Currently, hand geometry is employed at over 8,000 locations, including the Colombian legislature, San Francisco International Airport, day care centers, welfare agencies, hospitals and immigration facilities. The advantages of a palm print are similar to the benefits of a fingerprint in terms of reliability, although palm print readers take up more space. The most successful device, the Handkey, looks at both the top and side views of the hand using a built-in video camera and compression algorithms. Devices that look at other hand features are also under development by several companies, including BioMet Partners, Palmetrics, and BTG.

- Iris Patterns

  The advantage of iris scanners is that they do not require the user to focus on a target, because the pattern of flecks on the iris are on the eye's surface. In fact, a video image of the eye can be taken from up to three feet away, which allows for the use of iris scanners at ATM machines. In visually impaired persons with intact irises, the iris can still be captured and encoded with iris imaging products that have active iris capture (e.g., the ATM application). Since cataracts are a disease of the lens, which is behind the iris, cataracts do not affect iris scanning in any way.

- Retinal Patterns

  Retinal scans are performed by directing a low-intensity infrared light through the pupil to the blood vessel pattern on the back of the eye. Most uses of retinal scanners involve high-security access control, since they offer one of the lowest false rejection rates (FRR) and a nearly 0% false acceptance rate (FAR). However, since retinal imaging requires a clear view of the back of the eye, cataracts can negatively impact the retinal image quality.

- Facial Features

  Facial verification and recognition is one of the fastest growing sectors of the biometrics industry. Its appeal lies in the fact that it most closely resembles the way we as humans identify one another. Most commercial efforts have been stimulated by the fast rise in multimedia video technology that is placing more cameras in the home and workplace. However, most developers have had difficulty achieving high levels of performance. Nevertheless, specific

applications, such as screening welfare databases for duplicates and airport lounges for terrorists, are likely to appear in the future.

Comparison of Biometrics

Biometric technology is one area that no segment of the IT industry can afford to ignore. Security-wise, it benefits across the spectrum, from IT vendors to end users, and from security system developers to security system users. All these IT sectors must evaluate the costs and benefits of implementing such security measures. Different technologies may be appropriate for different applications, depending on perceived user profiles, the need to interface with other systems or databases, environments, and other application-specific parameters.

Table 2.1 The Large Number of Factors Involved In Biometrics Comparison

Characteristics	Ease of Use	Error Incidence	Accuracy	User Acceptance	Required Security Level	Long-term Stability
Fingerprints	High	Dryness, dirt, age	High	Medium	High	High
Hand Geometry	High	Hand injury, age	High	Medium	Medium	Medium
Retina	Low	Glasses	Very high	Medium	High	High
Iris	Medium	Poor lighting	Very high	Medium	Very high	High
Face	Medium	Lighting, age, glasses, hair	High	Medium	Medium	Medium
Signature	High	Changing signatures	High	Medium	Medium	Medium
Voice	High	Noise, colds, weather	High	High	Medium	Medium

Biometrics : Applications

The practical applications of biometric technologies are diverse and expanding, as new needs are identified. By and large, biometric applications fall into two main categories: law enforcement and civilian applications. The law enforcement community is perhaps the largest biometric user group. Police forces throughout the world use the technology to process criminal suspects, match finger images and bring guilty criminals to justice. Those biometric applications not involving crime detection utilize some form of access control. This will either involve the physical access of people to secure areas, or securing the access to sensitive data. In other word access control is either physical access or data access. Whether securing benefit systems from fraud, preventing illegal immigrants from entering a country, or prisoners from leaving a prison - controlling access is the underlying principle. Access control ensures that authorized individuals can gain access to a particular area and that unauthorized individuals cannot.

Some examples of biometric applications are listed below.

- Banking
- Computer Access
- Electronic Benefits Transfer (EBT)
- Immigration
- National Identity
- Physical Access
- Prisons
- Telecommunications
- Time & Attendance

Banking

Banks have been evaluating a range of biometric technologies for many years. Fraud and general breaches of security must be controlled if banks are to remain competitive in the financial services industry. Automated teller machines (ATMs) and transactions at the point of sale are particularly vulnerable to fraud and can be secured by biometrics. Other emerging markets include telephone banking and Internet banking, both of which demand the utmost security for bankers and customers alike.

Computer Access

Fraudulent access to computer systems affects private computer networks and the Internet in the same way: confidence is lost and the network is unable to perform at full capacity until the breach in security is patched. Biometric technologies are proving to be more than capable of securing computer networks. This market area has phenomenal potential, especially if the biometrics industry can migrate to large-scale Internet applications. As banking data, business intelligence, credit card numbers, medical information and other personal data becomes the target of attack, the opportunities for biometric vendors are rapidly escalating.

Electronic Benefits Transfer (EBT)

Benefits systems are particularly vulnerable to fraud. The battle against fraud has been waged by welfare departments across many states for years. A variety of technologies are being evaluated, although fingerprint scanning is particularly widespread. AFIS technology and one-to-one verification are used to ensure that

the benefit claimant legitimately receives a benefit check. Another development that may revolutionize the payment of benefits is Electronic Benefits Transfer (EBT), which involves loading funds onto a plastic card. The card can then be used to purchase food and other essentials in shops fitted with special point-of-sale smart card readers. Biometrics are well-placed to capitalize on this phenomenal market opportunity and vendors are building on the strong relationship currently enjoyed with the benefits community.

Immigration

Terrorism, drug-running, illegal immigration and an increasing throughput of legitimate travelers is putting a strain on immigration authorities around the world. It is essential for these authorities to quickly and automatically process law-abiding travelers and identify the lawbreakers. Biometrics are being employed in a number of diverse applications to make this possible. The U.S. Immigration and Naturalization Service (INS) is a major user and evaluator of biometric technologies. Systems are currently in place throughout the U.S. to automate the flow of legitimate travelers and deter illegal immigrants. Biometrics are also gaining widespread acceptance in Australia, Bermuda, Germany, Malaysia, and Taiwan.

National Identity

Biometrics are beginning to assist governments as they record population growth, identify citizens, and prevent fraud occurring during local and national elections. Often this involves storing a biometric template on a card which, in turn, acts as a national identity document. Fingerprint scanning is particularly

strong in this area and programs are already underway in Jamaica, Lebanon, the Philippines, and South Africa.

Physical Access

More and more organizations are using biometrics to secure the physical movement of people. Schools, nuclear power stations, military facilities, theme parks, hospitals, offices and supermarkets across the globe employ biometrics to minimize security threats. As security becomes more important for parents, employers, governments and other groups, biometrics will be seen as a more acceptable and therefore essential tool. The potential applications are endless. Biometrics could even be employed to protect our cars and homes.

Prisons

Prisons, as opposed to law enforcement, use biometrics not to catch criminals, but to make sure that they are securely detained. A surprising number of prisoners simply walk out of prison gates before they are officially released. A wide range of biometrics are now being employed worldwide to secure prison access, police detention areas, enforce home confinement orders, and regulate the movement of probationers and parolees.

Telecommunications

With the rapid growth of global communications, cellular telephones, dial inward system access (DISA), and a range of telecommunication services are under attack from fraudsters. Cellular companies are vulnerable to cloning (a new phone is created using stolen code numbers) and new subscription

fraud (a phone is obtained using a false identity). Meanwhile, DISA - which allows authorized individuals to contact a central exchange and make free calls - is being targeted by telephone hackers. Once again, biometrics are being called upon to defend this onslaught. Speaker ID is well suited to the telephone environment and is making inroads into these markets.

Time & Attendance

Recording and monitoring the movement of employees as they arrive at work, take breaks, and leave for the day was traditionally performed by "clocking-in" machines. However, manual systems can be circumvented by someone "punching in" for someone else, a process known as "buddy punching." This disrupts time management and unit costing exercises and costs companies millions of dollars. Replacing the manual process with biometrics prevents abuses of the system. In addition, biometrics can be incorporated with time management software to produce management accounting and personnel reports.

Keystroke Dynamics

Keystroke dynamics, also called typing rhythms, analyze the way a user types at a terminal by monitoring keyboard input 1,000 times a second. This is analogous to the early days of the telegraph, when users identified each other by "the fist of the sender." The advantage in the computer environment is that neither enrollment nor verification detracts from the regular work flow. Despite its appeal, however, efforts at commercial technology have failed.

Signatures

Static signature capture is becoming quite popular as a replacement for pen and paper signing in bank card, PC and delivery-service applications (e.g., Federal Express). Generally, verification devices use wired pens, pressure-sensitive tablets, or a combination of both. Devices using wired pens are less expensive and take up less room but are potentially less durable. To date, the financial community has been slow to adopt automated signature verification methods for credit cards and check applications because signatures are still too easily forged. This keeps signature verification from being integrated into high-level security applications.

2.2 The Study Of Voice Recognition

A Brief History

According to Chamberlain (1993), Bell Telephone in New Jersey produced the first working speech synthesizer that emitted speech sounds, called the 'Voder', in 1939. In 1947 a device called the Spectrogram was developed, providing the foundation for the voice recognition work that was to follow. In the 70's, template-based approaches were adopted. However, research and development into voice recognition technology in the 1970's was unsuccessful due to the speed and processing limitations of the available computers. Development of faster microprocessors in the 1980's saw more voice-dependent and voice-independent systems reaching the market. At this time, voice recognition products were developed for people with disabilities. The software increased their independence, helping them to participate more fully in society. The early 1990's witnessed many manufacturers launching their own voice recognition

systems. Dragon Systems, a company based in Newton, Massachusetts, was the largest manufacturer of voice-dependent systems in the United States. In 1993, Dragon and IBM were hard at work to develop a small vocabulary, Continuous Speech Series at the Massachusetts Institute of Technology. Since 1993, automatic voice recognition has come a long way and is approaching the science fiction capabilities of HAL, the infamous computer in 2001: A Space Odyssey. Today, the best dictation programs can handle continuous speech, recognize recurring word patterns, get smarter with use and have plenty of timesaving features. Rapid advances in the technology and fierce market competition have given consumers much more value for their money.

Voice Recognition : How It Work

First, to operate a computer through voice, the user must learn how to dictate in a word-by-word manner known as 'discrete speech'. In other words, the computer cannot recognize individual words if they are spoken the way people usually speak in fluent sentences or 'continuous speech'. Next, the user must 'teach' the system to recognize his or her voice through a combination of training and usage. We all pronounce individual words in different ways, and voice recognition software cannot simply recognize everyone's voice right off. As the user speaks to the system, the software creates a user-specific voice file that contains a lot of information about his or her voice qualities and pronunciations. The system uses this information to make its best guess at what each word is as it is dictated. The process of 'familiarizing' the voice recognition software with an individual voice takes time. When a user takes the time to properly train and use the voice recognition system, which creates a strong

and accurate voice file, the system will supply the correct word most of the time. However, the system will never achieve a 100% accuracy rate in all situations. Sometimes the computer just doesn't get it right and suggests the wrong word. The user must then stop and correct the system.

Voice recognition follows these steps:

1. Spoken words enter a microphone.
2. Audio is processed by the computer's sound card.
3. The software discriminates between lower-frequency vowels and higher-frequency consonants and compares the results with phonemes, the smallest building blocks of speech. The software then compares results to groups of phonemes, and then to actual words, determining the most likely match.
4. Contextual information is simultaneously processed in order to more accurately predict words that are most likely to be used next, such as the correct choice out of a selection of homonyms such as *merry*, *marry*, and *Mary*.
5. Selected words are arranged in the most probable sentence combinations.
6. The sentence is transferred to a word processing application .

According to Feldman (1996) the speech recognition engine converts the speech to a stream of text. The first step is to break the digitized sound into segments. These are matched against stored samples called 'phonemes,' which correspond to vowel and consonant sounds. The phonemes are then matched against the phoneme sequences that correspond to the words in the system's dictionary, and the message is converted

to a stream of text. The natural language processor extracts the meaning from the text based on stored grammar and the dialogue's context. The speech generation function is the inverse of speech recognition. With this function, a string of text is converted to a sequence of sounds based on a fixed vocabulary, similar to the one used for speech recognition, and a set of rules.

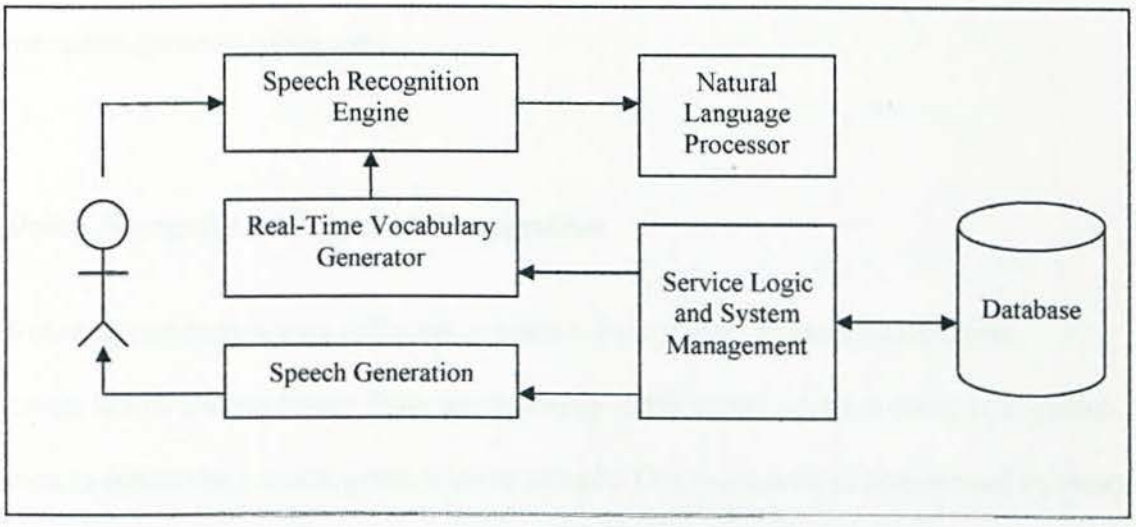


Figure 2.4 Components Of A Speech Processing System

Two types of speech recognition are in use today: voice-dependent and voice-independent recognition. Voice-dependent recognition can be considered the first generation of voice recognition. The computer is dependent on recognizing the user's voice and the vocabulary that the user recorded into the machine. Thus, the computer learns the individual's utterances and can replicate them by comparing them to a template copy. Voice-dependent recognition is based on matching a spoken word or phrase such as 'call mom' or 'read my mail' with a preprocessed speech sample, generally using the same speaker to optimize accuracy. The other type of voice

recognition is voice-independent recognition. This system recognizes speech components such as vowels and consonants and matches groups of these speech components with words in a dictionary. When voice-independent recognition is combined with a natural language processor, it is elevated to speech understanding, and the meaning of the spoken message can be extracted in a manner similar to the way humans respond to speech. The flexibility and natural dialogue possible with voice-independent recognition make it the technology of choice for call screening and menu navigation applications.

Voice Recognition Vs Speech Recognition

Voice recognition is very different to speech recognition. Speech recognition detect words and sentences from an incoming audio signal whereas voice recognition tries to detect the speaker given a voice sample. Our *voiceprint* is determined by many different factors: size of the vocal cavities (throat, oral, nasal) as well as the characteristics of the vocal cords themselves. Your voice is further modified by the way you speak - the way your mouth, lips, tongue, jaw and teeth move (these are called articulators). Therefore, the chances of two people having the same vocal characteristics are impossible.

2.4 Analysis On Existing Products

Most of the reviewed products are either biometrics based or voice recognition based systems. These products have given me an idea on how to approach the layout of my project and what the most suitable design would look like. Other than that, by evaluating them, it has enabled me to point out the pros and cons of the features displayed for each system.

The review has also given me a clearer picture of the system to be built. Hence, I hope that through evaluation made, I can proceed to developing a good login authentication system using voice recognition. I would present in this section a snapshot of the current products available in the market. For each system, I will describe first the companies involved in its development, followed by explanation on the system's architecture – that is the general layout of the system, then a list of their target users – people or industries benefiting, and end with the system's requirements (if provided). In hope to enable a clearer description of features for comparison, the evaluation done is portrayed through the pros and cons section. A pros and cons of features will be evaluated at the end of this each review section, thus this would display the features I would take into consideration in the development and design of my system.

[1] VoiceCop

http://www.voiceware.co.kr/

1. **Manufacturer :** VoiceCop Inc.

2. **Overview :**

VoiceCop is one of the biometrics security technologies - speaker verification, which verifies users by unique characteristics of the voices. It uses remote verification through internet and telephone network.

3. **System's Requirements :**

a) Windows 95/98/ME/NT 4.0/2000 Professional or Server

CPU : Intel Pentium 166MHz or higher

RAM: 10MB or higher

b) UNIX

OS: Sun Solaris (2.6, 2.7, 7), HP-UX, AIX, Iris, SCO UNIX, Linux

c) CTI Applications

User Verification for Phone Banking, Stock Inquiry

General Client/Server Application (User Authentication)

Home Automation: Voice Door/Safe Lock

4. **Unique Features :**

- **Easy Voice Password Registration**

With only 3 times voice password registration, VoiceCop can capture one's voice characteristics completely; thus it makes possible for the users to use

voice verification solution with no technical fear at all.

- **High Recognition Rate**

  Not only adapting conventional simple voice signal model comparison but also duplex voice model comparison using sentence-independent phoneme model which composed of 52 phonemes regardless of gender and age, VoiceCop has significantly improved authentication rate.

- **Intelligent Verification Engine - Speaker Adaptation**

  VoiceCop automatically reflects the periodic (7 to 10 days) changes of voice to its voice password database. As a result, it completely solved fatal verification rate decline problem as time goes by.

- **Perfect Voice Password**

  It can continuously monitors the differences of one's registered voices and unintentional registration errors. Eventually, VoiceCop maximizes the capability of building prefect voice password.

- **High Recognition Speed**

  The system builds and compares voice password database in a form of phoneme, not each character, so that any combination of voice password is available without missing. Moreover, the recognition speed of VoiceCop is greatly improved by such voice password combination architecture.

6. Target Users / Applications :

a) Home banking

b) Home trading

c) E-commerce and other commercial site user authorization

d) Door lock

e) Computer hardware log in system.

Pros:

- There are one free-online demo available. Unfortunately it's a text-to-speech module not a voice recognition module.
- This product convince me that speaker verification is more advantageous in technological and economical aspects compared to the other biometric security technologies since it requires no special hardware devices or has no danger of key loss and unauthorized use.
- It can run successfully on a Unix platform – a reassurance for me to try to run my project on a Linux (Unix-based Operating System)

Cons :

- No review on other products.
- No detailed explanation on the technical aspects of the system's architecture – particularly the matching techniques and the verification methods.

[2] Dragon NaturallySpeaking

http://www.transcriptiongear.com

1. **Manufacturer:** Dragon Dictate Professional

2. **Overview :**

The Dragon NaturallySpeaking family of products is the fast, easy and accurate way to turn speech into text. Users can dictate into virtually any windows-based application at speeds up to 160 words per minute and achieve higher levels of accuracy than ever before

3. **System's Architecture :**

Uses a speech recognition architecture, similar to what's been discussed in subsection 2.3.2., the system has four major modules – with each applies to different applications. There are as follows :

a) Dragon NaturallySpeaking Professional Solutions

A tool that enables corporate and professional users to work faster and more efficiently by speaking and words will appear on screen in letters, spreadsheets, and forms.

b)Dragon NaturallySpeaking Legal Solutions

A tool that enable legal professionals to create and format documents, enter time and billing information, send e-mail, and more - all by voice.

c) Dragon NaturallySpeaking Medical Solutions

A tool that enable medical professionals to dictate speech directly into electronic medical records, create patient notes, fax referral letters, and work on the Web

d) Dragon NaturallySpeaking Public Safety Solutions

A tool that enable public safety professionals to create investigation reports, interview summaries, and other documents, fill-in forms, send e-mail; and more - all using voice.

4. Specialized Application Support

a) Internet Explorer 5, 6

b) AOL® 6, 7, 8

c) Word 97, 2000, XP

d) WordPerfect® 10

e) Excel 97, 2000, XP

f) Outlook® Express 5, 6

g) Outlook® 97, 98, 2000, XP

h) PowerPoint® 97, 2000, XP

i) Create multiple vocabularies using text documents

5. Unique Features :

- Speeds up routine tasks on the PC, such as creating documents, entering data, launching applications, sending e-mail, completing forms, and browsing the Web.
- Reduces or eliminates transcription processing costs and delays for all professional organizations, including healthcare, legal and government.
- Increases productivity when away from the office by transcribing recorded dictation from mobile devices, including Microsoft Pocket PC and commercial digital recorders.

- Promotes prompt, standardized reporting and documentation procedures.
- Helps to protect employees from Repetitive Strain Injuries (RSI).
- Is available with Enterprise IT tools for management of multi-seat corporate installations.
- Is available with a full set of APIs for developers and system integrators who want to fully speech enable any workflow application.

6. Target Users :

a) Corporate and professional users

b) Legal professionals

c) Medical professionals

d) Public safety professionals

e) Businesses and professionals in the insurance, finance, research, and manufacturing industries

f) Disabled PC users looking to optimize productivity with their PCs

g) Home users that want to have fun with their computers.

Pros :

Some of the advantages of analyzing this system :

- Different modules work for different applications
- Short set-up and user enrollment
- Dictate into most Microsoft Windows-based applications
- Control menus and dialog boxes in most Microsoft Windows XP & Windows 2000-based applications by voice

- Simultaneous dictation and command modes are all formatted and edited by voice
- Mouse control can be done by voice – a feature that benefit people with disabilities.
- Support USB audio
- Natural Punctuation is very important during user enrollments
- Dictation Playback – a very convenient method to confirm user's enrollment.
- Save audio with text dictation
- It uses a handheld digital recorder
- Dictate into Pocket PC
- Third-Party Correction – applicable to any open source – based applications.

Cons :

- No explanation on system's requirements.
- No documentation on concept of the technical aspects of the various modules of the system.
- No free demos provided – interested readers have to buy via on-line first

[3] Vocent Solution Inc.

http://www.vocent.com

1. **Manufacturers :** Vocent Solutions Inc.
2. **Partners :**

Vocent partners fall into three categories, each having specific areas of concentration:

Technology And Platform - Voice recognition, voiceprint authentication, telephony, hardware

Nuance http://www.nuance.com/

SpeechWorks http://www.speechworks.com/

Dialogic http://www.dialogic.com/

NMS Communications http://www.nmscommunications.com/

Channel Access http://www.channelaccess.com/

BeVocal http://cafe.bevocal.com/

Help Desk - Service management and web-based password reset automation

Remedy http://www.remedy.com/.

Courion http://www.courion.com/

Security - Single sign-on, password consolidation, provisioning, authorization server, authentication server

Netegrity http://www.netegrity.com/.

Courion http://www.courion.com/

BMC Software Inc. http://www.bmc.com/.

M-Tech http://www.psynch.com/.

3. Overview :

Vocent Solutions Inc. is the leader in voiceprint solutions that securely automate the critical and costly function of authenticating callers. Headquartered in Mountain View, Calif., the privately held company applies industry standards to develop platform-independent solutions that combine proven voice technology with unique application-level functionality to maximize caller automation processes while maintaining comprehensive security.

4. System's Architecture :

The system is divided into two major parts.

(i) Vocent's Voice Secure Password Reset

This software makes a digital representation of its 5,000 employees' network passwords by recording a voice sample for each customer or employee - typically the sound of that person pronouncing the numbers zero through nine. That information is then stored and used for future comparisons.

(ii) Vocent's Voice Secure Confirmed Caller

The identity of subsequent callers is verified by asking each caller to pronounce a randomly selected sequence of those digits into the phone receiver. This software uses sophisticated algorithms and speech recognition technology to match the caller to the recorded digital representation of that person's voice. The random number sequence guards against the use of recorded voices to trick the system.

Figure 2.5 Vocent Password Reset

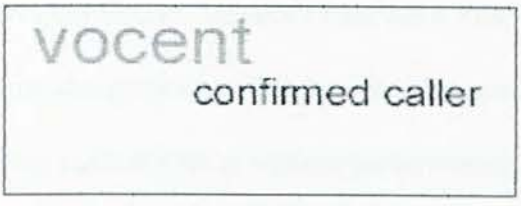


Figure 2.6 Vocent Confirmed Caller

5. Unique Features :

- They have the vision to extend voiceprint authentication into IP-based data networks, such as the Internet, to give enterprises a way to confidently identify consumers beyond today's availability of voiceprint authentication in telephone applications.
- Provides packaged solutions for telephony-based voiceprint authentication; namely, Vocent™ – Confirmed Caller and Vocent™ – Password Reset, which can automate authentication in more than 90-percent of telephone-based transactions that otherwise require expensive agent time to confirm a caller's claimed identity or reset a password.

- Voiceprint authentication provides an ideal and practical way to combat identity fraud and assure privacy.
- Vocent's Password Reset 2.0 also includes a new Multi-System Module that reduces complexity in resetting passwords for users with multiple accounts or who access multiple backend systems.
- Vocent's Password Reset 2.0 resolves this issue through a speech recognition dialog which saves users and help desk agents time by eliminating the need to individually reset passwords for each backend system associated with a specific account.
- In addition to these new features, Vocent's Password Reset 2.0 includes improved application manageability. The new release provides additional reports for monitoring authentication system performance, such as average daily enrollments, number of first-time enrollments and the percentage of enrollments that required multiple attempts.
- Vocent's Password Reset 2.0 is a robust, full-featured enterprise software solution for automating password resets

6. Target Users / Applications:

Biometric security in the areas of e-commerce, mobile commerce and risk management.

Pros :

- It's recognition and verification simultaneously – an aspect I would like to consider for this project.
- Future solutions by Vocent will leverage voiceprints across the data network. To provide a strategic opportunity for enterprises and consumers to secure automatic transactions regardless of the channel used.
- Vocent solutions incorporate a biometric approach to securely authenticate users attempting to gain access to an account.
- Rather than using only passwords, pin numbers or other knowledge-based content that can be compromised, voiceprints provide strong authentication by combining 'what you know' with 'who you are'.
- This way, even if imposters find out personal information; e.g., account ID, social security number, or mother's maiden name, a voiceprint comparison would prevent them from gaining access to on-line financial or health-related information, or purchasing an item using a stored payment card.
- This technology can be used to obviate the need for the 20+ passwords a typical person must remember.

Cons :

- Even though the description touches most of the technical aspects I want to know, unfortunately it left out the hardware and software requirements.
- There are no free online demos available for users to test the system.

[4] VoiceProtect

http://www.voice-security.com

1. **Manufacturer :**

   Voice Security System Inc.

2. **Overview :**

   Founded in 1999, Voice Security Systems Inc. has been involved in the development of voice/speaker verification/authorization systems using embedded technologies.

   A decade ago, securing access to devices, doors, services, etc. using speaker verification technology was still in its infancy. Processors were very slow when compared to today's modern electronics. Storage space was also at a premium. Many systems required prohibitively large storage facilities for verification software and voice template storage, even for a modest number of users. To resolve these problems, a method was developed which could adequately verify an individual's identity using their voice input, that did not require a large database of enrollment templates or sophisticated voice processing equipment.

   The evolution of this technology is **Voice Protect®**.

   [Quoted from the website]

3. **System Architecture :**

This new breakthrough allows speaker-verification to be burned onto an existing microprocessor within a device. Examples of usage of this technology are cell phones (to eliminate cell phone fraud), ATM manufacturers (to eliminate pin # fraud) and automobile manufacturers (to dramatically reduce theft and carjacking).

This method is the only stand-alone technology that does not require large user database management, thus protecting the privacy of the user and shier biometric data.

The software, algorithms and templates can be stored on the microprocessor that a device already employs to operate the functions of the electronic hardware inside.

Voice Security Systems Inc.'s approach to fraud prevention is currently the simplest, most cost-effective and reliable Biometric security method currently available. Obtaining a voice template only requires a simple microphone or telephone. Because the Intellectual Property can be hardware-based, it does not require expensive equipment to enable it to operate, it does not require extensive redesigning for each application, nor does it require any costly administrative or maintenance expenses to keep it operational. The only costs are one-time, non-recurring engineering event.

This breakthrough technology also enables speaker verification features to be easily added to many existing microprocessor based devices such as PDAs, cell phones, computers, and smart cards. Most of these devices will not require the addition of any dedicated voice processing hardware or additional storage

to utilize the **Voice Protect®** method of speaker verification (a user's complete set of enrollment data fits in less than 800 bytes).

4. **Unique Features :**

- Non-intrusive enrollment (takes less than 30 seconds).

  Voice Protect has a very small storage requirement for enrollment data, which is easily stored on secure physical media such as smart cards, iButtons, or a consumer device such as a cell phone. The enrollment data is all that is needed, other than a live voice sample, to verify the claimed identity of the user.

- When verification is performed, there is no spoken or test prompt as to what the correct pass phrase is, making it virtually impossible for a thief to impostor.

  Even if the impostor knows the correct pass phrase, the **Voice Protect** method accurately rejects them based on the biometric features unique to the enrolled user.

- The user can change their pass phrase more frequently to increase security (very much like text passwords are handled today).
- The **Voice Protect** verification method adjusts for small physical changes in the user's voice over time, allowing long term use of enrollment data. In fact, the level of verification confidence can actually improve within the first several uses of the system.

- Methods of speaker verification can effectively operate without the benefit of an external server or hard drive. The entire software program and templates can be operated and stored in the protected device itself.
- Optionally, the voice templates used for validation can be loaded at verification time from a smart-card or stored in external RAM due to their small size.
- They believe they have a smaller, more robust and more efficient process of speaker authentication than anyone else currently in the industry. The storage requirement for the "voice print" which is compared to a live sample is a little as 780 bytes (this is the actual requirement for the keypad demonstration on this site).
- The code needed to perform the validation is very compact and will fit into many inexpensive chips available today; including smart cards. For example, adding the verification process to the Keypad Demonstration program only increased the program size by 2K.
- Those involved in adding speaker verification to the security layer of their application, are only now beginning to realize the benefits of a simple, unobtrusive user enrollment and verification process such as theirs. Their method the first to operate without the need of a cohort set of enrollment data, requiring only three samples of a phrase chosen by the user. They were also the first to streamline enrollment procedure from several hours or even days, in some cases, down to only a few seconds. Creation of a new "voice key" (enrollment) can be completed in under 20 seconds. The process can run on as little as an 8-bit processor running at 8 MHz.

6. System Performance

Here is a summary of test results using version 2.0.18 of the keypad test control available at this site:

Total Users Tested : 133

Total Users Enrolled : 77

Cohort Testing - Enrolled (77) vs. All (133 - one attempt each due to limited data)

Total False Accept (known phrase) : 94 / 10164 or < 0.93%

Verify Testing - Enrolled (66) vs. All (133 - 40 attempts each)

Total False Accept (unknown phrase) : 364 / 421652 or < 0.09%

- 66 of the users recordings were unable to generate satisfactory enrollment scores due to CO-articulation differences between the digits recorded.
- These user's enrollments were invalid, thus they were omitted from the Cohort Testing (where the user tries to impostor 'knowing' the secret pass phrase).
- Data from ALL one-hundred thirty-three users was used for the verification testing to determine the "real-world" probability of someone trying to gain access to a stolen device for which the pass-phrase is unknown.
- The verify test phrases were all different from each other and different from the "correct" phrase that was compared to the impostor attempt.

- All of the phrases used consisted of "combination-lock" sequences; three two-digit numbers spoken as: "62...76...53".
- In an actual application, the user can create a "voice key" by saying virtually any phrase in any language as long as it meets a minimum energy requirement (about 2 seconds of speech). We believe this would further decrease the possibility of someone guessing the correct phrase, although the verification method still rejects most closely matched speaker(s) even if they know the phrase.
- These numbers will vary depending on the user, phrase selected and specific hardware used to extract the voice features.

5. **Target Users / Industries :**

   Opportunities present for Voice Security Systems Inc. are, the Banking industry, the personal computer industry and the Department of Defense.

Pros :

- Explaining was done using simple English
- They do explain the features of the product adequately touching almost all aspects of said system.
- Provide general knowledge on biometric technology, authentication and network security which is a good advantage for project's foundation knowledge.
- The system architecture was carefully explained – differ the site from the others.

- Storage of the user's voice data and the hardware and software required for the speaker verification process and fit into a standard double gang switch box (5"x5"x3").
- The inclusion of system's performance (previously mentioned) has given me a thorough look on what to expect during this project's testing. The idea of having a 2 seconds voice input (the enrollment recording) is motivated by this system's success.
- This system is very valuable in order to help aide my project's development.

Cons :

- There is a free online demo for this system , which is available from the website , but unfortunately, the link was broken.

[5] Voice Reset

http://www.sentrycom_files\voicereset_main.htm

1. **Manufacturer :** SentryCom

2. **A Brief History :**

   SentryCom's mission is to deliver reliable, cost-effective and easy to use authentication solutions to secure assets within the extended enterprise. The company designs, develops and markets biometric voice authentication

solutions to increase and enhance security while improving end-users privacy and confidence.

3. System's Architecture :

VoiceReset takes advantage of the fact that the best solutions never require a change in behavior. Users will continue to call in order to reset their passwords, however the call will now be automated, self-serving and secure. The only hardware needed for VoiceReset is the telephone, which is already in place and does not require additional investments or training for the users.

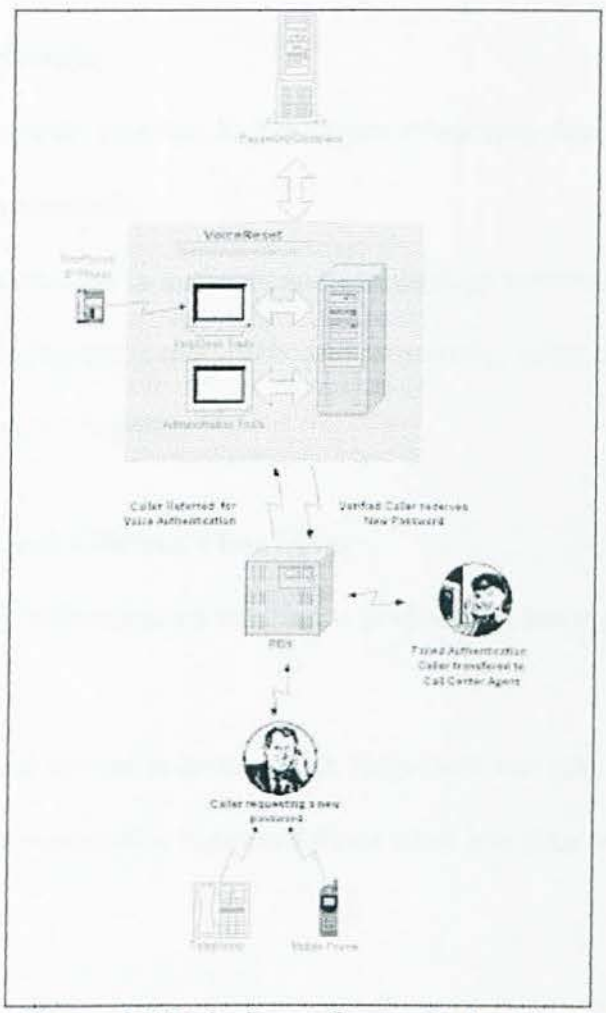


Figure 2.7 VoiceReset System's Architecture

4. **System's Requirements :**

Other than the usual workstation this system requires a telephone.

5. **Unique Features :**

- **Reduced Cost and Improved Productivity**
  - i) Manual resets are recurring cost
  - ii) Voice authentication based solution is one-time investment
  - iii) Reduces the requirements for Help Desk staffing
  - iv) Help Desk representatives can handle more complex tasks
  - v) Increased productivity by reducing time spent on hold to Help Desk

- **Tightened Security**
  - i) Biometrics provides higher degree of security than traditional PINs and passwords
  - ii) Biometric voice authentication technology minimizes identity theft
  - iii) Unlike tokens, cards, PINs, and passwords, voice cannot be shared, stolen, or forgotten

- **Convenient and Efficient Time Saver**
  - i) Voice authentication solution is available 24 hours a day, 7 days a week
  - ii) Typical manual reset through a Help Desk can take up to 15 minutes
  - iii) Voice verification Password Reset takes less than 60 seconds

6. **Target Users / Applications :**

   VoiceReset can be used externally for resetting Web access passwords, and internally for Legacy system Access, business administrators and end users.

Pros :

- Explanation on system's architecture is completed with a figure that show's the entire automated control process of voice authentication.
- To ensure the highest level of security, a user's identity is verified using caller phone ID, voiceprint authentication and a PIN number .
- It is not a login system but nonetheless it has been a good reference for my project .

Cons :

- No free online demonstration for user to test.
- The system's requirements is poorly explained.
- A user must have basics in authentication to really understand the methods used in implementing this system.

[6] RECOVC

http://www.alphaworks.com

1. **Manufacturer :** IBM ViaVoice
2. **Overview :**

The Recognition-Compatible Voice Coder (RECOVC) is a distributed speech recognition software that uses low-bit-rate voice coder/decoder (CODEC). It is designed to work in applications in which compressed speech is to be processed either by an Automatic Speech Recognition (ASR) engine for text-to-speech conversion or by a Speaker Recognition/Verification engine. Unlike other commonly used low-bit-rate CODECs, which cause degradation in recognition rates, this product compresses speech while keeping the recognition rates completely intact. The compression is based on extraction and coding of the acoustic features used by the recognition engine. A technique enables the reconstruction of a good-quality speech from the acoustic features.

3. System Architecture :

The demonstration application enables distributed speech recognition and remote playback. At the client, the voice is captured and compressed into a low-bit-rate stream (8 Kbps) by the RECOVC encoder. Audio capturing and real-time transmission via RTP are carried out using JMF (Java Media Framework).

At the server, the recognition features are decompressed from the low-bit-rate stream and large-vocabulary, continuous speech recognition is carried out by IBM's ViaVoice engine (Version 8.0). The recognized text is transmitted back to the client using a TCP socket and presented to the user. In parallel, the voice signal is reconstructed by the RECOVC decoder and can be played back at both the server and the client.

4. **System Requirements :**

Operating System : Windows NT or 2000 (for both client and server)

Server Hardware : Pentium 300 MHz or above

96 MB of RAM

Server Software : IBM ViaVoice for Windows, Release 8

5. **Unique Features :**

- Speech recording and recognition are performed either at distant remote locations or at different points in time, and where the transmission channel bandwidth or storage space is limited

6. **Target User / Applications :**

RECOVC is intended for use in distributed speech recognition (DSR) systems, where the accessing of voice portals on the Web and other Interactive Voice Response (IVR) services from Internet phones, cellular phones, or other portable devices, as well as from Personal Digital Assistant (PDA) devices that record voices to be recognized later, are examples of such applications.

Pros :

- How this system work is briefly but efficiently described
- The techniques used are not familiar to me but nevertheless it opens a new opportunity to learn new approach to speech recognition.

Cons :

- There were no figures or tables that can help summarize the system's concept.
- No free-online demos to try
- Users are recommended to have background in signal processing for easier understanding.

There are many more products in existence out there. The reason why only these six systems were chosen is not based on popularity but because the knowledge they provide is crucial in my attempt to continue my study and research on the architectural aspects of the systems previously mentioned.

Next few sections are studies done on several knowledge domains crucial in developing this project at hand.

2.5 The Study of AI Methods

2.5.1 Artificial Neural Network (ANN)

Artificial neural network takes their names from the network of nerve cells in the brain. It provides a unique computing architecture. Recently ANN has been found to be an important technique for classification and optimization problem. ANN is capable of performing non- linear mapping between the input and output space due to its large parallel interconnection between different layers and the non- linear processing characteristics. An artificial neuron basically consists of a computing

element that performs the weighted sum of the input signal and the connecting weight. The sum is added with the bias or threshold and the resultant signal is then passed through a non-linear element type. Each neuron is associated with three parameters that can be adjusted during learning; these are connecting weights, the bias and the slope of the non-linear function.

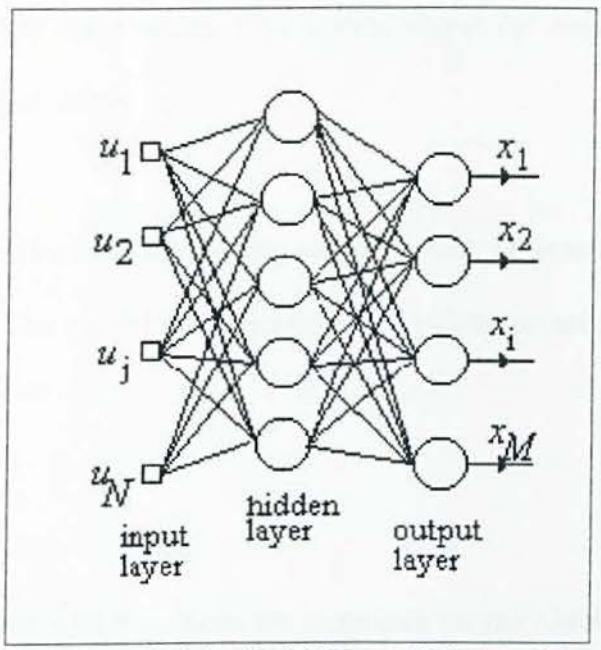


Figure 2.8 Artificial Neural Network

Learning Of ANN

The learning of the ANN may be supervised in the presence of the desired signal or it may be unsupervised when the desired signal is not accessible. Rumelhart developed the back propagation algorithm, which is central to much work on supervised learning in multi layer Neural Net. A feed forward structure with input, output, hidden layers and non linear sigmoid functions are used in this type of network. In recent years many different types of learning algorithm using the incremental back propagation algorithm, evolutionary learning using the nearest neighbor MLP and a fast learning algorithm based on the layer by layer optimization procedure are suggested. In case of

unsupervised learning the input vectors are classified into different clusters such that elements of a cluster are similar to each other in the same sense. This method is called competitive learning, because during the learning process a set of hidden units compete with each other to become active and perform the weight change. The winning unit increases its weights on those links with high input values and decreases them on those with low input values. This process allows the winning unit to be selective to some input values.

The simplified ANN (supervised) training and prediction process can be illustrated by the following steps. The crucial pre-processing and validation are discussed separately.

Stage One:

Collect the training set, which includes the input data for the ANN to "see" and the known target data for ANN to learn to output. For stock price predictions, for example, the training set and target data would naturally be historical stock prices. A vector of 100 consecutive historical stock prices, for instance, can constitute training data and with the 101st stock price as a target datum.

Stage Two:

Feed the input data to ANN; compare ANN output with the known target, and adjust ANN's internal parameters (weights and biases) so that ANN output and the known target are close to one another—more precisely, so that a certain error function is minimized.

Step Three:

Feed ANN some future input data (not seen by ANN); if ANN is well trained and if the input data are predictable, then ANN will give accurate predictions.

2.5.2 Learning Vector Quantization (LVQ)

Vector Quantization is a technique used for compression of speech and image data. The basic idea is to represent the input vectors with a smaller set of prototypes that provide a good approximation to the input space, where the vectors constitute the input space(will be explored in detail in subsection 2.6.6).

Learning Vector Quantization(LVQ) was developed by Kohonen. This is a supervised learning technique that can classify input vectors based on vector quantization. It is a type of competitive networks (Figure 2.8), where output units compete for the right to respond. The goal here is to provide a way to do *clustering* (Figure 2.9) - divide the data into a number of clusters such that the inputs in the same cluster are in some sense similar. Clusters (also called *classes*) are predefined and a set of data is labeled. The aim is to determine a set of prototypes that best represent each cluster.

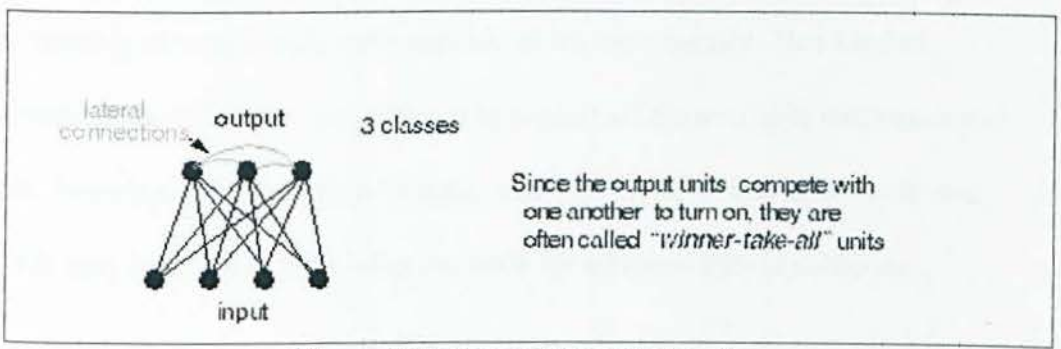


Figure 2.9 Competitive Network

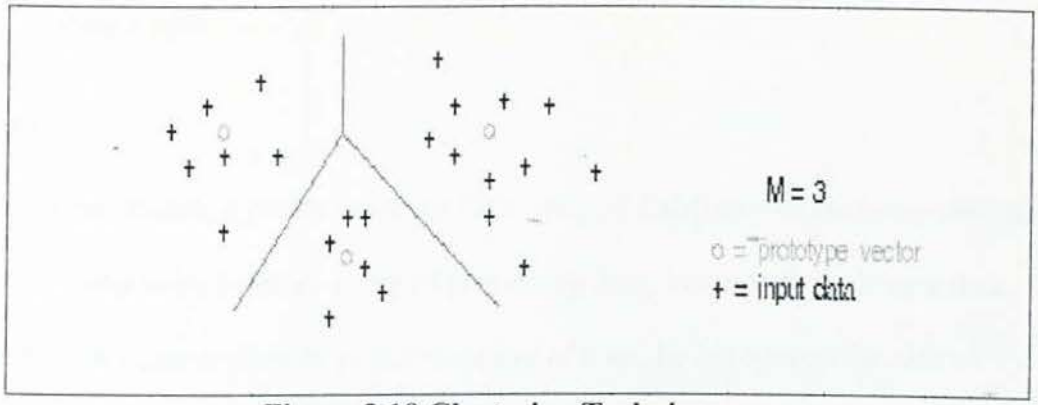


Figure 2.10 Clustering Technique

Why LVQ?

The minimization of classification errors is the main objective in most pattern recognition applications. This is often approached by modeling the probability densities of the competing classes, but it is often not possible to assume any proper parametric density model, the lowest error rate is obtained by concentrating on the actual discrimination between the classes.

The methods based on neural networks may outperform other methods in tough problems, where the prior knowledge cannot help much in the classification and the system characteristics must be learned automatically from the data. For those complicated situations it is advantageous that the algorithm consists of a large number of very simple units capable of learning locally. This kind of structure can be efficiently parallelized to exploit all the available computational power. Simple algorithms often have the tendency to be widely applicable and provide easy integration with other methods for efficient hybrid solutions.

2.5.3 *Fuzzy Logic*

History

In 1964, Lotfi Zadeh, a professor of the University of California of Berkeley was the first to develop fuzzy logic as a way of processing data; instead of requiring a data element to be either a member or non-member of a set, he introduced the idea of partial set membership. Actually the idea of Fuzzy Logic was derived from the Boolean Logic. Instead of using numerical values, fuzzy logic uses general terms.

(e.g. If a person is 1.83 metres of tallness, they are 'rather tall'.)

Definitions

1. **Fuzzy Logic**

   Fuzzy Logic is a powerful problem-solving methodology with wide application in embedded control and information processing.

2. **Fuzzy Set Theory**

   - Classes or groupings of data with boundaries that are not sharply defined.
   - Is a set whose elements have degrees of membership.

3. **Traditional or Crisp Set**

   A crisp set contains elements which belong to that set. If an element does not fully belong, then it is not included. Every element in the universe of discourse can be assigned a value representing its membership in particular sets. When dealing with crisp sets, this value is either zero or one; either an element belongs, or it does not.

Table 2.2 Classical Logic vs. Fuzzy Logic

No.	Classical Logic	Fuzzy Logic
1	Deep understanding of a system	Incorporates an alternative way of thinking
2	Exact equation	Allow modeling of complex system using our own knowledge and experience
3	Precise numerical values	Allows knowledge representations using subjective concepts

Real-life Applications

- **Air conditioning:** gradually slows down the cooling system as the room temperature approaches the desired setting.
- **Cruise control**: determines ambient acceleration or deceleration and controls the countering application of gas and brake.
- **Ship boilers:** monitors the temperature, pressure, and chemical content to ensure stability.
- **Video cameras:** identifies when the subject of a video shot is moving and when motion is caused by the cameraman's vibrations.
- **Washing machines:** optimizes the wash cycle by examining the load size, fabric mix, and quantity of detergent.

Advantages and Disadvantages

Advantages:

a. Ability to describe systems linguistically through rule statements.

b. Mimic human decision making to handle vague concepts.

c. Rapid computation due to intrinsic parallel processing nature.

d. Ability to deal with imprecise or imperfect information.

e. Resolving conflicts by collaboration, propagation and aggregation.

f. Improved knowledge representation and uncertainty reasoning.

g. Modeling of complex, non-linear problems.

h. Natural language processing/programming capability.

i. Alternative design methodology which is simpler and faster.

- Reduces the design development cycle.
- Design complexity is simplified.
- Time to market improved.

j. Better alternative solution to Non-Linear Control System

- Higher control performance
- Simpler implementation
- Hardware costs reduced

Disadvantages:

a. Threatens the integrity of scientific thoughts.

b. Expands the possibilities of things like computer programming.

c. Lack of a formal design methodology.

d. Resulting system is not analytical, and therefore any mathematical analysis on paper is impossible with current methods.

e. Highly abstract and heuristics.

f. Needs experts for rule discovery (data relationships)

g. Lack of self-organizing & self-tuning mechanisms of Neural Network

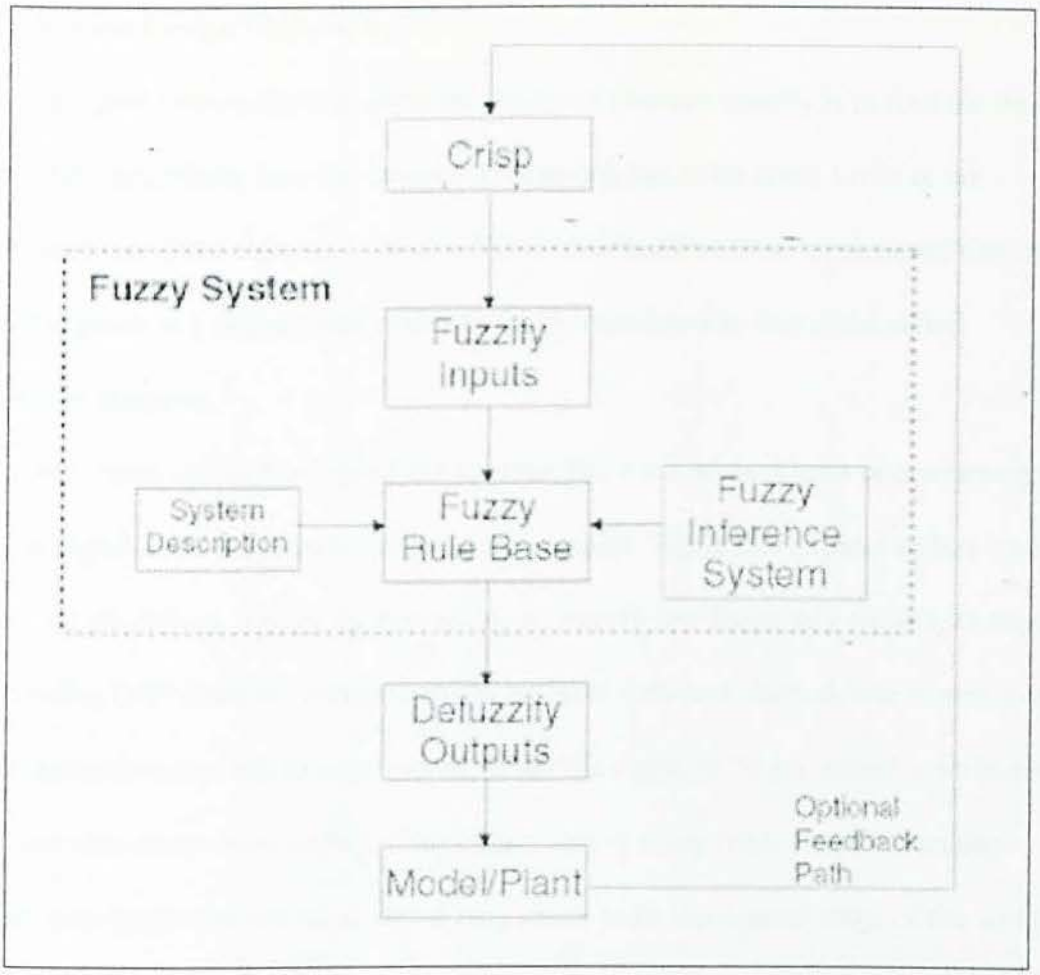


Figure 2.11 Basic Block Diagram Of A Fuzzy System

2.6 Digital Signal Processing (DSP)

Digital Signal Processing is distinguished from other areas in computer science by the unique type of data it uses: *signal*s. In most cases, these signals originate as sensory data from the real world: seismic vibrations, visual images, sound waves, and others. DSP is the mathematics, the algorithms, and the techniques used to manipulate these signals after they have been converted into a digital form. This includes a wide variety of goals, such as: enhancement of visual images, recognition and generation of speech, compression of data for storage and transmission.

DSP System Design Philosophy

When a signal processing task turns up, the basic concern usually is to find the best algorithm prescribing how the required processing has to be done. Little or no attention more often than not is paid to the representation of the input signal assuming that it is given in a digital form or can be easily converted to that if the signal originally is analog.

However based on the mentioned assumption, there are no problems in converting the analog signal at the input into its digital counterpart. While that is more or less true under the conditions typical for processing relatively low frequency signals, in more demanding DSP cases the situation might be quite different. Indeed, when extremely wide dynamic range has to be achieved, when the signal to be processed is wideband and contains components at high frequencies and in many other cases the analog-digital conversions of the input signal may prove to be the crucial stage of the whole signal processing process.

The most common of DSP operations are signal sampling and quantizing, which is the basis of any analog-digital conversion. Little could be done in an attempt to adapt them to the specific conditions of a given signal processing case. Actually it is only possible to vary the time intervals between signal sample taking instants at sampling and to compare the precision of the sample value rounding-off at quantizing.

The limitations in matching the common analog-digital conversion process to the specifics of the signal processing task at hand would prevent obtaining good results.

Significance of the sampling and quantization processes

To process signals digitally, they obviously have to be presented in the appropriate digital format. Therefore the original analog signal, before processing, has to be converted into a digital one, it has to be digitized. Often there are no problems. Then digitizing and subsequent processing can be performed with sufficiently good results by using the available Analog-Digital Converters (ADC) and the standard DSP techniques.

In other words, it is absolutely crucial that correct DSP-defined signal digitization procedures are used for conversion of the original analog signal into an appropriate digital format. Good signal processing results can be expected only if this condition is met. Once a signal is digitized, the features of the obtained digital signal, good as well as bad, are fixed and nothing can be done to change them. In an ideal world, these features would exclusively depend and actually would copy the features of the original analog signal. The reality is different. The two basic operations of any analog-digital conversion, namely, sampling and quantization, impact the characteristics of the digital signal substantially. The characteristics of the analog signal at the ADC input and of the digital signal at its output are just similar rather than identical. How large and significant the differences between them are depend on the digitizing methods and their implementations applied.

Thus the significance of the sampling and quantization operations is determined by the fact that many essential digital signal characteristics, impacting the whole signal processing process substantially, depend on the methods and techniques used to perform them.

The only way how to reduce this often undesirable dependence is increasing the sampling frequency and/or decreasing the quantization step-size. However these possibilities then are poor and limited. In addition, this approach in many cases produces an increased number of bits requiring more complicated hardware for the subsequent processing of the obtained in this way digital signals.

DSP advantages

Real-life world is basically analog. So are the signals reflecting the processes going on in that world. Computers, on the other hand, are digital. Therefore there is a gap between the real world and computers. This gap is filled by DSP that provides for analog signal conversion into a digital format and for preprocessing such digital signals.

2.6.2 Audio Sampling

For the most part, the mechanisms of the natural world around us, including sound, operate in the analog domain. And so the transducers used to convert sound into electrical signals (microphones) and vice-versa (loudspeakers) are analog as well. A microphone produces an electrical signal with an infinite number of amplitudes which can be amplified to a suitable level for further processing such as mixing, recording, transmission and reproduction. The chain of devices that amplify, process and carry the electrical signal from the source (the microphone) to the destination (the loudspeaker) — and the wires that connect them — can be thought of as a medium. Unfortunately, the devices in this medium introduce inherent, undesirable

impairments (linear distortions, nonlinear distortions and noise) that degrade the quality of the signal. The impairments contributed by each device are additive — that is, they accumulate. Thus, the number and individual performance of the devices in the medium determine its overall performance. This puts a limit on the number of devices through which an analog audio signal can pass before the impairments become unacceptable.

However, it is possible to eliminate many analog signal-handling difficulties by digitizing the electrical signal before sending it through the medium. Digital audio systems convert the original analog signal to a binary digital signal which has two well-defined states: zero and one. Undesirable electrical impairments affect the digital signal just as they affect the analog signal, but they have no effect on the information the digital signal carries as long as the device receiving the signal determines that the binary signal levels are within the threshold values for the "zero" and "one" states. Such systems restrict message distortion to the analog-to-digital (A/D) and digital-to-analog (D/A) conversion processes, thereby improving the transparency of the medium. The medium remains transparent as long as it maintains a certain level of signal-to-noise ratio (SNR), beyond which the "cliff effect" occurs and the transmission shuts off. This article examines some of the basic audio analog-to-digital conversion concepts, emphasizing the sampling process.

Sampling considerations

Sampling is the first step towards digitizing audio signals. It consists of measuring the amplitude of the analog audio waveform at periodic intervals, T. The main concern is to represent the original analog values with adequate precision. The measurement accuracy depends on the frequency at which the audio signal is measured, or sampled.

The sampling frequency must be at least twice (preferably more than twice) the highest audio frequency being sampled.The sampling process consists of multiplying the analog audio signal with a stream of repetitive pulses — a pulse amplitude modulation (PAM) process.

Quantizing considerations

The next step in analog-to-digital conversion is quantization. In this process, the samples are assigned a binary number approximating their sampled value. Quantizing divides up the sampled voltage range into 2^n-1 quantizing intervals, where "n" is the number of bits per sample (the sampling resolution). For example, an 8-bit system can identify 2^8 (256) discrete sampled signal values (255 quantizing intervals). The amplitude of such a signal can occupy the entire quantizing range. However, low-amplitude audio signals would be quantized with considerably fewer discrete levels, resulting in significant quantizing errors. These quantizing errors are correlated with the signal and perceived as distortion. With higher-level signals, the quantizing errors are uncorrelated with the signal and perceived as random noise. One can reduce quantizing errors by increasing the number of bits per sample, increasing the sampling frequency (oversampling), or both.

2.6.3 Cepstrum Method

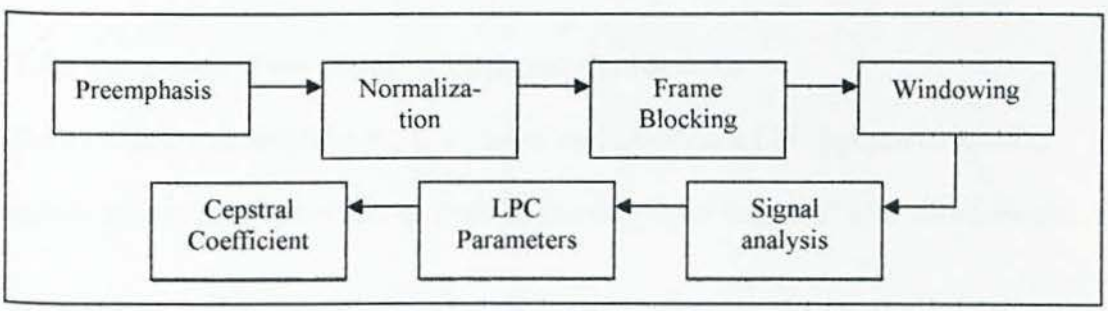


Figure 2.12 Block Diagram of A speech recognition system

1. Pre-emphasis

The speech signal is filtered with a first-order FIR filter to spectrally flatten the signal. One of the most widely used preemphasis filter of the form is normally used.

2. Normalization

After preemphasis, each word has it's energy normalized. Based on the energy distribution along the temporal axis, it is computed the center of gravity, and this information is used as reference for temporal alignment of the words

3.Frame Blocking

The preemphasized speech signal, is blocked into frames of N samples, with adjacent frames being separated by M samples, with N and M represent a set of numbers.

4.Windowing

Each individual frame is windowed to minimize the signal discontinuities at the borders of each frame.

5.LPC Parameters

The next processing step is the Linear Predictive Coding (LPC) analysis using the autocorrelation method of order.

LPC Parameter Conversion to Cepstral Coefficients

The LPC cepstral coefficients, c_m, are a very important LPC parameter used in speech recognition. They can be derived directly from the set of LPC coefficients

The cepstral coefficients, which are the coefficients of the Fourier transform representation of the log magnitude of the spectrum, have been shown to be more robust for speech recognition than the LPC coefficients

7. Cepstral Distance

The cepstral coefficients provide an efficient computation of the log-spectral distance of two frames. For LPC models that represent smoothed envelopes of the speech spectra, it is usually used a truncated number of cepstral coefficients.

2.6.4 Vector Quantization (VQ)

Vector Quantization VQ is the process where a continuous signal is approximated to a digital representation (quantization) considering a set of parameters to model a complete data pattern (vector). The field of VQ is strongly related to the reduction of the number of bits necessary to convert analog signals considering distortion or fidelity criterion. Vector quantization techniques are highly used in speech technology applications for efficient coding feature and pattern selection .

Scalar And Vector Quantization

Scalar Quantization quantizes one sample at a time.

Vector Quantization quantizes a block of samples at a time. This may be:

- A vector of samples (one dimensional), typically for speech and audio coding.
- A block or matrix of samples (two dimensional), typically for image coding.
- Exploits correlations between samples.

- Rate-distortion theory shows that a lower bit rate may be obtained with the same distortion when using vectors rather than scalars (or, a lower distortion at the same bit rate as scalar quantization).

Basic Concept

Encoder and decoder have a vector codebook or just "codebook" of representative source vectors. Encoder buffers a block of samples to be coded. Searches for the best match in the codebook (usually minimum least-square error). Encoder transmits the index of the best-match vector to the decoder. Decoder uses codebook as a lookup table to reconstruct the signal.

Encoding and decoding are thus asymmetrical: encoder has to do a search of the codebook; decoder simply a lookup. For code vectors of dimension there are subtract-and-square operations per vector. This is repeated for the code vectors in the codebook. Requires fast DSP's (digital signal processor) to perform this search. This is termed an "exhaustive search" as all candidate code vectors are searched.

VQ Applications

- For image coding, either mean-removed (mean+shape+gain) VQ used (scalar quantization for mean and gain, vector quantization for shape).
- For transform-based image coding, VQ of the DCT-transformed samples is possible. For speech coding, synthesis filters are used. Essentially, VQ is used to encode the parameters of a filter block (input signal and filter coefficients), which is then used "synthesize" the block of speech at the receiver.
- Full treatment of VQ for speech & images is very involved.

2.7 The Study of Project Softwares

2.7.1 Matlab

Matlab is a commercial "Matrix Laboratory" package which operates as an interactive programming environment. It is a mainstay of the Mathematics Department software lineup and is also available for PC's and Macintoshes and may be found on the CIRCA VAXes. Matlab is well adapted to numerical experiments since the underlying algorithms for Matlab's built in functions and supplied m-files are based on the standard libraries LINPACK and EISPACK.

Matlab program and script files always have filenames ending with ".m"; the programming language is exceptionally straightforward since almost every data object is assumed to be an array. Graphical output is available to supplement numerical results.

Online help is available from the Matlab prompt (a double arrow), both generally (listing all available commands):

>> help

[a long list of help topics follows]

and for specific commands:

>> help fft

[a help message on the fft function follows].

Paper documentation is on the document shelf in compact black books and locally generated tutorials are available and are used in courses.

How to quit Matlab ?

The answer to the most popular question concerning any program is this: leave a Matlab session by typing

quit

or by typing

exit

to the Matlab prompt.

Batch jobs

Matlab is most often used interactively, but "batch" or "background" jobs can be performed as well. Commands can be debugged interactively and stored in a file (`script.m', for example). To start a background session from an input file and to put the output and error messages into another file (`script.out', for example), this line is entered at the system prompt:

nice matlab < script.m >& script.out &

Other work or logout can be done at the machine while Matlab grinds out the program. Here's an explanation of the sequence of commands above.

1. The "nice" command lowers matlab's priority so that interactive users have first crack at the CPU. This must be done for non-interactive Matlab sessions because of the load that number-crunching puts on the CPU.
2. The "< script.m" means that input is to be read from the file script.m.
3. The ">& script.out" is an instruction to send program output and error output to the file script.out. (It is important to include the first ampersand (&) so that

error messages are sent to the file rather than to the screen - omit the ampersand and the error messages may turn up on *other* people's screens

4. Finally, the concluding ampersand (&) puts the whole job into background.

Unique Features of Matlab

- Matlab has many types of matrices which are built into the system. A 7 by 7 matrix with random entries is produced by typing

  rand(7)

- To generate random matrices of other sizes and get help on the rand command within matlab:

  rand(2,5)

  help rand

- Matlab has built-in variables like pi, eps, and ans. Their values can be learned from the Matlab interpreter.

Programming In Matlab

MATLAB is also a programming language. By creating a file with the extension .m you can easily write and run programs. If you were to create a program file myfile.m in the MATLAB language, then you can make the command myfile from MATLAB and it will run like any other MATLAB function. You do not need to compile the program since MATLAB is an interpretative (not compiled) language. Such a file is called an m-file.

2.7.2 Linux

Linux was originally created by Linus Torvalds of the University of Helsinki in Finland. Linus based linux on a small PC-based implementation of UNIX called *minix*. Near the end of 1991, Linux was first made public.

Advantages of Linux

- Full multitasking – multiple tasks can be accomplished and multiple devices can be accessed at the same time.
- Virtual memory – Linux can use a portion of the hard drive as virtual memory which increases the efficiency of your system by keeping active processes in RAM and placing less frequently used or inactive portions of memory on disk. Virtual memory also utilizes all system's memory and doesn't allow memory segmentation to occur.
- The X Window System – a graphic system for UNIX machines. A powerful interface
- That supports many applications and is the standard interface for the industry.
- Shared Libraries – each application shares a common library of subroutines it can call at runtime. This saves a lot of hard drive space on the system.
- Non-proprietary source code – The Linux kernel use no code from AT&T or any other proprietary source. Other organizations, such as commercial companies. The GNU project. Hackers. And programmers from all over the wotld have developed software for Linux.

2.11 Chapter 2 Summary

The literature review is a critical look at the existing research that is significant to the work being carried out. This review would describe, summarize, evaluate and clarify literature. Hence in the literature review I have done, I have evaluated and gave critical appraisal to related existing products viewed via the Internet. In doing so I have come up with a view of what I want to feature on my login system. I have also pin pointed features with drawbacks needed to be overlooked, and also features with advantages that can be used.

In doing this review, I was given the opportunity to learn the other systems, thus I came up with a conclusion of what required of a developer to build a voice authentication system that can satisfy user's security needs.

Besides the products, I had also reviewed on the software and technology used in voice application development, which I had summarized in this chapter.

Chapter 3, next, is all about the methodology chosen for this project.

CHAPTER 3

-METHODOLOGY-

3.1 Overview of Methodology

The word 'methodology' is a combination of two Yunanian words – 'methodos' and 'logos'. Methodos means a way or manner, and logos means knowledge on how to do analysis. In short, methodology is further defined as a collection of procedures, techniques, tools and documentations. It helps software developers to build a system according to plan and produce a high quality product.

There are a lot of methodology models in existence currently. Ranging from the classic life cycle models to the innovative evolutionary models – they provide adequate analysis on project duration, budget and requirements to software developers. Below are some of the most popular methodology models :

- Waterfall Model
- Waterfall Model With Prototype
- V Model
- Transformation Model
- Prototyping Model
- Operational Specification Model

3.2 The Prototyping Model

A system development model gives a standardized and systematic approach to the project development. Software prototyping is an information system development methodology based on building and using a model of a system for designing, implementing, testing and installing the system

The Prototype Modeling methodology has been chosen to aide in the development of this project because of its many advantages and the main reason this methodology was chosen is because of its iterative approach in modeling processes whereupon each prototype developed in each modeling stage is revised continuously and flaws detected during revision period will be corrected earlier. A prototype is the usually full-scale and functional form of a software design constructed.

The characteristics of prototyping are:

- Prototyping is based on building a model of a system to be developed
- Prototyping uses the model for designing the system
- Prototyping uses the model to implement the system
- Prototyping uses the model to perform both the system and the acceptance testing of the system.

In implementing the prototyping methodology in this system development there are a few steps to be taken. These steps are listed next :

- The system requirements are defined in detail, whereupon requirements are gathered through various discussions with my supervisor and fellow colleagues and also through reviewing a number of products available in the market representing aspects of an example of system proposed.
- A preliminary design is created for the new system.
- A first prototype of the new system is constructed from the preliminary design. Basically, the first prototype is a scaled down system, which represents an estimate of the characteristics of the final system.
- The first prototype has to be thoroughly evaluated, by noting its strengths and weaknesses. Remarks on system performance will be collected and evaluated.
- Thus first prototype will be modified based on the remarks, and second prototype is built.
- Second prototype is analyzed in the same manner as the first prototype
- Preceding steps are iterations, which are finally concluded when system has function fully satisfying system's own requirements and the users. Final prototype represents the final product desired.

Final system is thoroughly evaluated and tested. Routine maintenance is carried out on a continuous basis to prevent extensive failures and minimize downtime.

This methodology proposes several advantages like:

- It reduces development costs.
- Decreases communication problems.
- Lowers operations costs.

- Minimizes time required during maintenance phase.
- A system that meets users definitive needs is created.
- Reduces and saves manpower

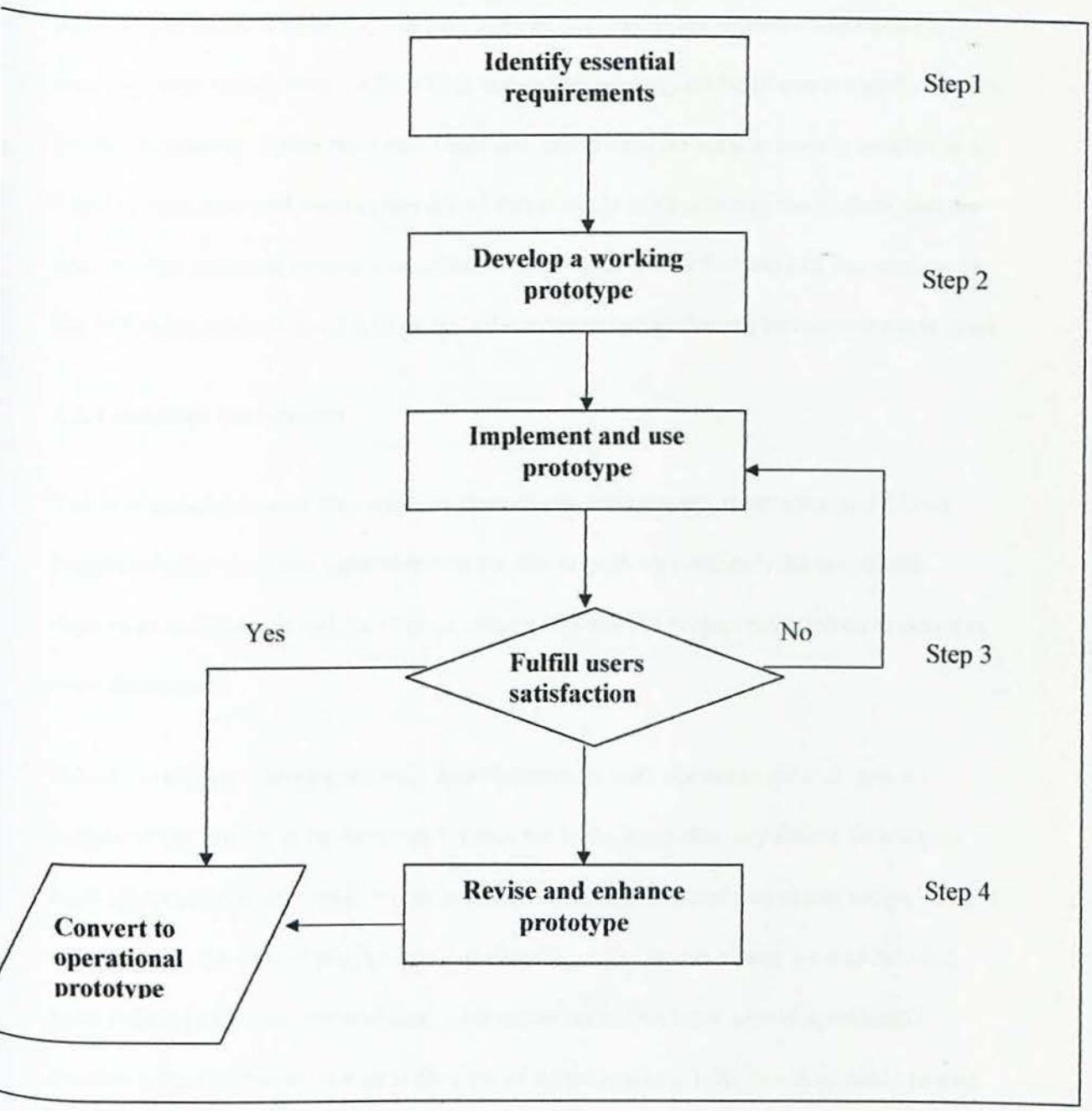


Figure 3.1 Prototyping Model

3.3 Information Gathering Techniques

In collecting information for this project, I had applied a few information-gathering techniques. Information and sources gathered were mainly concerning building a voice-based applications. Therefore, I had researched adequately on software available for creating voice recognition applications, and the technology on hand and currently used in biometric security. Aside from that I had also researched on what is usually needed in a login system, how and what approach AI can provide in developing the system, and the accessibility issues of system's interface to allow ease of use for users of this system. In the following subsections, I will explain the information gathering techniques made used.

3.3.1 Informal Discussion

The first technique used was informal discussions with project supervisor and fellow project developers. From these discussions, the project was defined, the scope and objectives ascertained, and the overall picture of what the project is expected to perform were determined.

Initially my fellow developers and I had discussions with our supervisor to gain an insight on the project to be developed. Once we had gained this, my fellow developers and I discussed among ourselves on how we would redefine our respective scope, so as to narrow down the overall project into our specific tasks. This is where we had decided upon defining our separate modules, whereupon each developer would specifically develop a module based on a specific type of AI techniques. Informal discussion paired up with book and source references lead to the determination of the technology to be used.

Aside from that, informal discussions has also lead to concluding on the design of interface, also what features and functionalities would be able on this system. The VIVA presentation had concluded with my moderator and supervisor giving ideas on changes to be made on the system. Hence informal discussions played a major role in aiding information gathering.

3.3.2 Internet Surfing.

The main source of information in analyzing available and similar systems was by surfing the Internet. It was easy to gain access to the Internet as the labs in faculty provided this facility. As is known, the Internet provides excessive information on any topic. Therefore by searching for my related field of study, I had gained plenty of resources on these types of information:

i) Digital Signal Processing

ii) Voice Feature Extraction

iii) Voice Recognition

iv) Artificial Neural Network

v) Fuzzy Logic

3.3.3 Reference Books

There are numerous books on Digital Signal Processing, Artificial Neural Network, Fuzzy Logic, Matlab, Red Hat Linux – that provide tools and techniques related to the development of the system. As an Artificial Intelligent majoring student, I have the advantage of owing some of the books, even borrowing them from our Main Library is never a hard task. They provide me an opportunity to learn up these subjects as they give a deeper insight on various concepts crucial to develop this system. The previous semester saw me taking up courses on Fuzzy logic and Neural Network, thus I had plenty references and adequate knowledge on them. These books had provided me the foundation for designing and developing this system.

3.3.4 Past Research

By reading past and similar researches done by previous seniors, I had an inkling and general idea of how to complete the first part of the thesis, which is the first stage of documentation from chapters 1 through 4. Aside from that I was able to compare the methodology and functionalities of previous systems, thus I had an idea of what functionalities I would like my system to perform. Nevertheless I could not find projects which developed the same exact system . But then there were numerous other voice recognition-based systems and these projects provided me an idea of how my project would be developed.

3.4 Chapter 3 Summary

The methodology chosen as a step-by-step guideline for the system development is the prototyping modeling. This methodology was chosen because of its iterative nature and requirements revision capabilities. As new developers of a software system, I felt that the methodology would provide me with clearer understanding of users requirements as they are revised constantly, and that would also save cost on maintenance.

The different types of information gathering techniques utilized were informal discussions with project supervisor, peers and colleagues, research of related reference material, review of past researches, surfing the Internet for information and last but not least conducting reviews on voice products and biometric systems available online. The last method has proved very useful in eliciting user requirements.

CHAPTER 4

-SYSTEM ANALYSIS & DESIGN-

4.1 Overview of Systems Analysis

Designing a system requires system analysis as the initial procedure before actually attempting to design the system. This procedure consists of a few major steps which aide in system development. Through analyzing current systems in the Literature Review, problems of current system are spotted and defined which leads to generating the problem statement that has to be solved in this system. Next, arises the determination of system requirements, which are collected by gathering significant information. Requirements gathered leads to the development of alternative solutions and finally choosing the most appropriate solution. Thus the primary deliverables from system analysis are the listing of system requirements.

The main purpose of this phase can be concluded as listed below:

- To learn how a similar system functions.
- To resolve system requirements by collecting user needs and by identifying major components to be included in the system
- To ensure the software development methodology proposed suitable for analyzing and developing system
- To determine hardware and software specifications to be used

4.1.1 *Functional Requirements*

Login ID

Each user will be ask to register for an account name (an ID) before each voice enrollment and verification.

- **Voice Recording**

  This function is crucial whenever a user uses the system. This is where user do voice recordings – three times for first-time enrollee and once to verify.

- **Playback Recording**

  This is an enhancement for the system where a playback is featured to aide user to confirm their own voice password.

  It is hope this function can help remind users of their own voice passwords.

- **Verification Status – Binary output**

  System will only generate a 'Yes' or 'No' answer to show user's verification / authentication, rather than outputting percentage or probability of the match.

4.1.2 *Non-Functional Requirement*

- Feasibility and Reliability of User Interface

All buttons and links must function properly to avoid errors during authentication.

- User's are free from any nasal and throat diseases like cold, flu and also sore-throats
- User must be an English literate because this login system will opt English as the medium.
- System should be able to give security comfort to anyone using it regardless of age and sex.
- Response time should not exceed 10 seconds
- Quiet Environment – to help reduce white noise during voice recordings.
  It is important in order to get the raw feature of a voice that signifies an individual.

4.1.3 *Software Requirements*

- Matlab Version 6.5
- Red Hat Linux Version 7.2
- X Windows
- FLEXlm 8.0d, installed by the MATLAB installer
- Netscape Navigator 4.0 and above or Microsoft Internet Explorer 4.0 and above
- Windows 98

4.1.4 *Hardware Requirements*

- Intel 386 through Pentium III
- 320MB of hard Disc Space in character mode or 450MB with X Window
- 128 MB RAM - 256 MB RAM
- CD-ROM drive
- 8MB – 16MB of Memory
- CD-ROM drive
- 3.5-inch disk drive
- SCSI or IDE CD-ROM drive
- Microphone
- Speaker

4.2 *Overview of System Design*

This chapter of the thesis will touch on the process that changes user requirements into an application that will be developed conceptually or logically. This phase will select and plan a system that meets user requirements, hence enabling the development of the desired system. In designing a system, the output would be a complete design specification that describes features of the system, the components and elements of the system and the system appearance to the user. There are three main stages in the system design process:

i) Component Design

ii) Technical Design

iii) User Interface Design

4.3 Component Design

As the system incorporates voice enrollment and verification technologies (as illustrated in Figure 4.1), it will have a microphone, attached to a workstation (computer) using a cable. Other components of this voice authentication system consist of a processing algorithm and matching algorithm. These are usually hardware or software components, but I rather have it implemented as software algorithms running on the workstation.

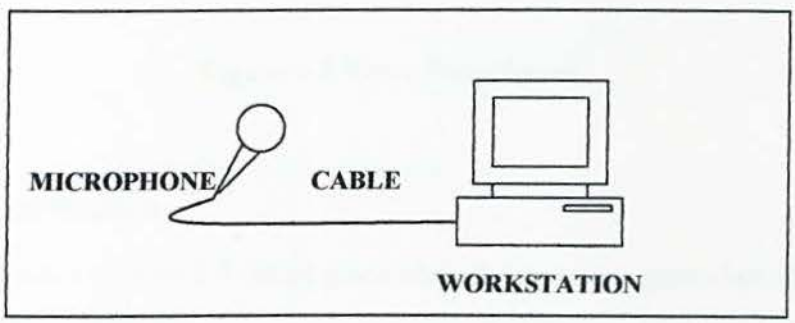


Figure 4.1 System Components

4.4 Technical Design

4.4.1 Voice Enrollment

The user enrollment process (Figure 4.2) consists of taking a sample of the user's voice. The sample taken is usually processed using a proprietary processing algorithm. The processing algorithm extracts specific information from the voice and stores this information in a data object called a 'voice template'. The voice template is used for

comparisons in user verification. The template cannot be used to recreate the original voice signal, but instead the live recording is fed through the algorithm to create another digital template which then can be compared to the original template.

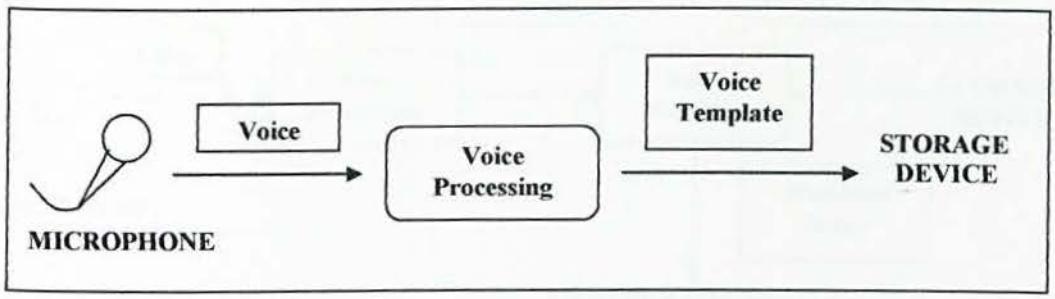


Figure 4.2 Voice Enrollment

4.4.2 Voice Verification

The user verification (Figure 4.3) takes place when the system requires voice authentication for each login process. A live recording of the user's voice is taken, processed into a voice template, and compared against the template taken during the enrollment process. The comparison is performed by a 'matching algorithm'. The matching algorithm typically gives a measurement of the degree of certainty of which the user was verified (e.g. 0 to 100 where 100 is an exact match and 0 is none at all). The measurement of certainty is also referred to as the 'matching score' (also called 'recognition score'). Since most voice matching takes place from voice samples that are taken at different points in time, no two voice signals can be identical, and consequently no match is ever an exact match. The system then has to determine what level it will accept as a match. The higher the level, the more false rejects the system will experience

(i.e. legitimate users that are rejected because a poor sample was taken). The system would need a threshold set on a per user basis to allow certain users with poor biometric qualities to pass at a lower level without sacrificing system security.

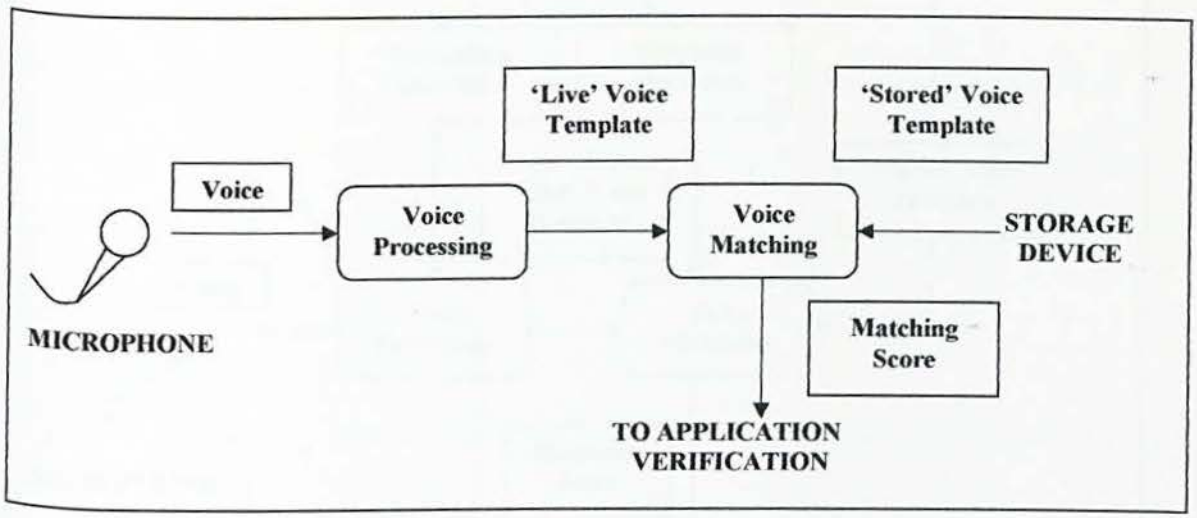


Figure 4.3 Voice Verification

4.4.3 Voice Authentication

The last two sub-sections explained the processes of enrollment and verification of user one by one. They cover both concepts separately and adequately. Since this project is about voice authentication where both concepts must combine in order to have a fully function system, this sub-section will guide you to the detailed design of the system, as shown in the figure below.

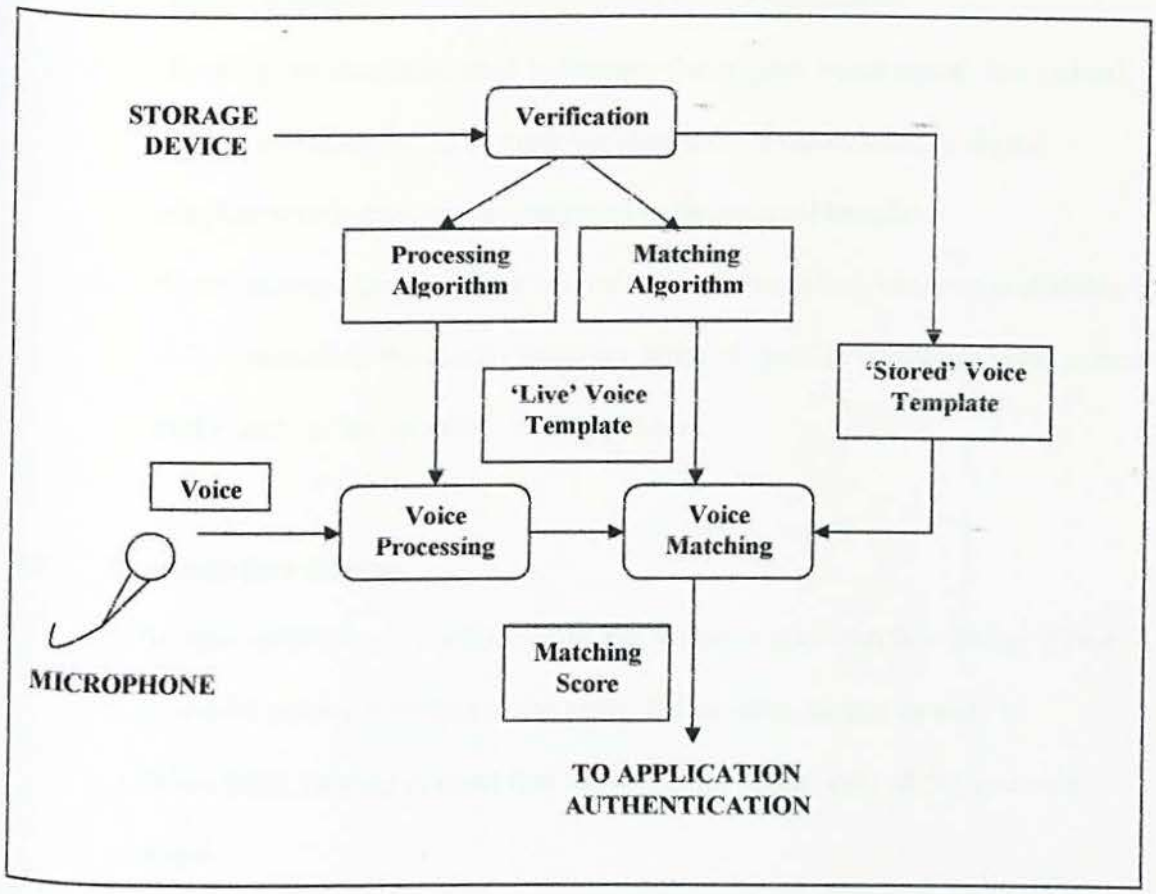


Figure 4.4 Voice Authentication

What happen during the whole login authentication process using voice recognition is briefly explain as follow:

i) Take sample of user's voice

ii) Sample taken is processed using processing algorithm

iii) The processing algorithm extracts specific info from voice and stores info as templates (also called 'stored template')

iv) Stored template is used for comparisons in user verification

v) The template cannot be used to recreate the original voice signal, but instead the live recording is fed through the algorithm to create another digital template which then can be compared to the original template.

vi) The matching algorithm then will calculate the matching score or probability before outputting the results as binary score. A 'yes' to verify and grant access to the user , a 'no' to show verifying failure.

4.5 *User Interface Design*

In the next subsections , I will describe the system's user interface design. Since there will be no nice interface to the users, I'll be using simple graphic to visualize them, bearing in mind that this is not the actual look of the system's interface.

4.5.1 *Main Menu*

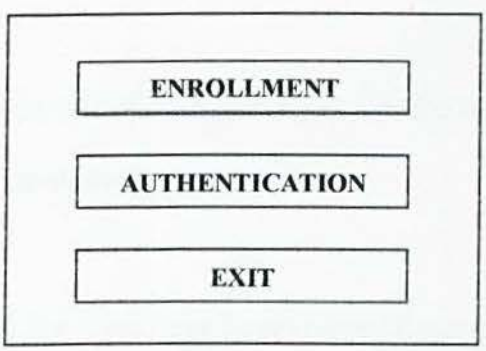


Figure 4.5 Main Menu Interface

This is the first thing that every user will see There are only three link buttons, 'enroll' , 'authentify' and 'exit'. From this main menu page, the user can gain access to the enrollment interface to enroll their voice , or to the authentication interface to verify

themselves , or even to 'exit' if there is nothing that the person would want to do at the moment.

4.5.2 *Enrollment*

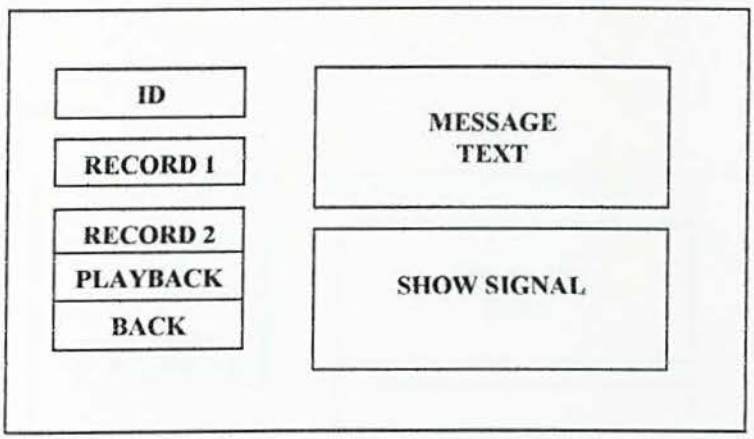


Figure 4.6 Enrollment Interface

- If user choose the 'enroll' button in the main menu , this interface will appear with several buttons with their own specific function ;

- The box at the top right side of the figure represents the ID that all user need to register before each enrollment.

- On the left hand side , there are three boxes named Record 1, Record2 and Playback. Record 1 and Record 2 represents the two recordings of voice from the user to enroll into the system. User need to click the button to speak through the microphone provided.

- The playback button will replay the user's voice after each enrollment.

- The exit button is to exit the interface back to the main menu when user's done with their enrollments.

4.5.3 *Authentication*

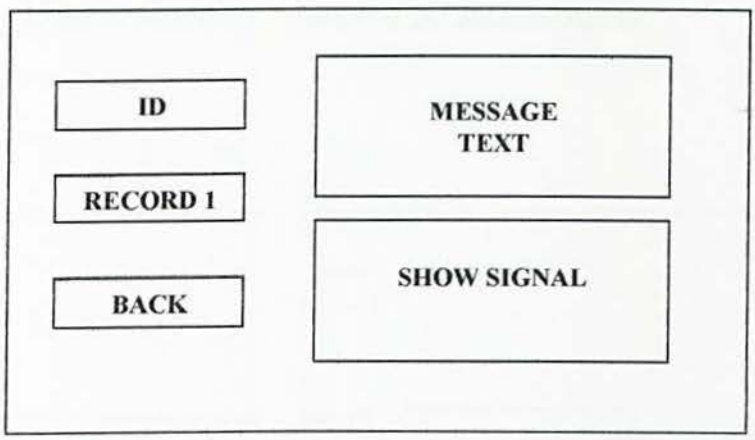


Figure 4.7 Verification Interface

- If the user choose the 'authentication' button in the main interface, this window will appear.
- As with the enrollment interface, user need to register / input the same ID used by typing it in the ID login box..
- But this time instead of having two recordings, the user need only to record his / her voice once using the record button that functions exactly like the record buttons in the enrollment interface.

- The message box will display certain messages according to the flow of both the enrollment and the authentication phases.
- The signal box will display signals recorded by user.
- An exit button is to exit the interface back to the main page.

4.6 System Flow

Next is a summary of the system flow – shown in the diagram below:

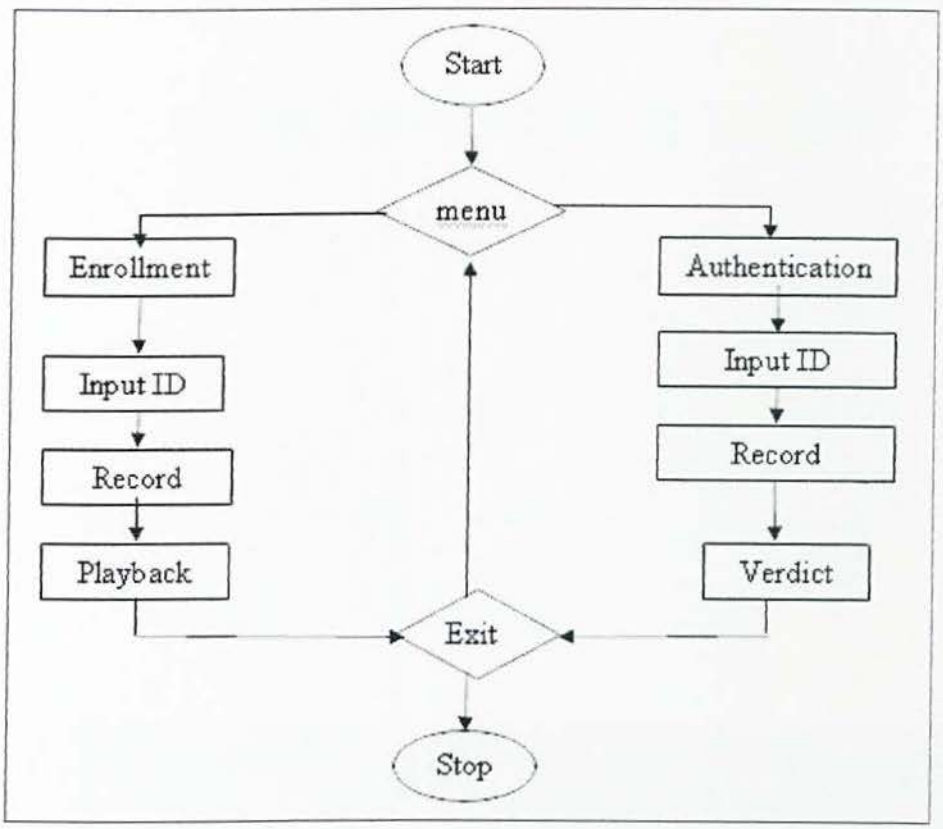


Figure 4.8 System Flow Diagram

Chapter 4 Summary

This chapter explains the modeling of requirements into processes of the system. The logical design of the system is then built using the process modeled. This whole system is designed using the top-down approach whereupon the system is decomposed from the main module into several modules, from high-level design to the low level design. The high level design gives a general description of the system, which is then divided into more detailed description in the low-level design

CHAPTER 5

-SYSTEM IMPLEMENTATION-

5.1 Overview of System Implementation.

System implementation is a process of developing a system based on the given requirements. Thus application tools and programming languages that are suitable for the development of the proposed system are chosen carefully.

5.2 Development Environment.

The environment chosen in which the system will be developed is the most important area of concern as with appropriate platform, software and hardware components than only will development process run without much hindrance. Generally the hardware and software components and tools used would speed up the development process.

5.2.1 Platform

Matlab was chosen as the platform for ease of implementation. The system functionality is available using Matlab 5 but Matlab 6 is required for the GUI. The Matlab Data Acquisition Toolbox (DAQ) is needed to perform the program recording.

In the proposal, it was stated that this system would be able to run on the Linux (RedHat Linux 7.2) operating system. In the early stages of implementation, a computer was successfully setup and reconfigured using this open-source based OS. Unfortunately, due to several un-avoided circumstances, it was impossible to get the installer to Matlab for Linux – software and a medium by which Matlab needs to run on the operating system itself, hence the system can only be implied on Windows.

5.2.2 Program Usage

The section covers the function calls for using the program from the Matlab command line. Please see Appendix B for GUI screenshots. The main function is to access the modules of the system which is enrollment, threshold creation, and authentication. The calling syntax is mainmenu. Available options from the main menu are User Enrollment and User Authentication.

5.2.3 High Level Organization

There is nearly a one to one correspondence between the design modules and implemented functions. Basic functionality from the design was implemented into the core functions accessible through the Matlab command-line. See Appendix C for files created and the calling structure.

5.3 Development of Proposed System

The following subsections will explain the development of this system login authentication using voice recognition. It will focus on the analysis of every speech technology used and also explain several functional details

5.3.1 Voice Acquisition

The first in the implementation phase is to get the voice signal. This is the process where voices are obtained in the form of analog inputs. Enrollments are done by the user by verbally speaking or saying his or her password twice.

The first is to create the codebook and the second is to create the user's threshold (will be further explained in subsection 5.3.5). The calling function is init_*sound.m*. This function initializes microphone input for user's to record their voice. All recorded voice is returned and stored as wave file (.WAV). Init_sound.m is called once each time a voice input is recorded, which takes place during both user enrollments and authentications.

5.3.2 Voice Processing

5.3.2.1 Analog to Digital Conversion

The analogue speech signal has been converted to a digital speech signal so that it can be discretely manipulated by a digital computer. The signal is sampled, quantized and coded. Following the rule of the Nyquist sampling theorem, which state that the analogue signal should be sampled at more than twice the highest operating frequency to avoid aliasing, the sampling rate was fixed at 16 kHz. The system uses the default quantization levels provided by the platform Matlab which is 16-bit or 2^{16} quantizing intervals. All coding were made by assigning unique binary numbers to each 16-bit quantization level

5.3.2.2 Enhancement

The recorded speech can be contaminated by background noise or interfering speakers. The transmission channel also deforms the speech. To help curb this problem the channel is set to be mono instead of stereo. A snippet of the coding is retrieved from file *init_sound.m* and is shown below.

```
ai = analoginput('winsound');
addchannel(ai, 1);      % 1 indicates mono; 2 indicates stereo
```

5.3.2.3 Endpoint Detection

The implementation for the login authentication system further addresses the issue of finding the endpoints of speech in a waveform. The code which executes the algorithm can be found in the file *locatespeech.m*. The algorithm removes any DC offset in the signal. If there was no silence removal, part of the voiceprint would come from the silence and an impostor could easily pass by just recording silence. If the DC offset is not removed, the zero-crossing rate of noise cannot be found in order to eliminate it from the signal. The algorithm finds the start and end of speech in a given waveform, allowing the signal to be analyzed.

The endpoint detection algorithm functions as follows:

1. Using the function removedc.m , the algorithm removes any DC offset in the signal. This is a very important step because the zero-crossing rate of the signal is calculated and plays a role in determining where unvoiced sections of speech exist. If the DC offset is not removed, the system will be unable to find the zero-crossing rate of noise in order to eliminate it from the signal.
2. Using the function avgmag.m and zerocrossing.m the average magnitude and zero-crossing rate of the signal as well as the average magnitude and zero-crossing rate of background noise was computed. The average magnitude and zero-crossing rate of the noise is taken from the first hundred milliseconds of the signal. The means and standard deviations of both the average magnitude and zero-crossing rate of noise are calculated, enabling the system to determine thresholds for each to separate the actual speech signal from the background noise.
3. At the beginning of the signal, the first point was searched where the signal magnitude exceeds the previously set threshold for the average magnitude
4. From this point, the search moved backward until the magnitude drops below a lower magnitude threshold.
5. From here, the previous twenty-five frames of the signal were searched to locate if and when a point exists where the zero-crossing rate drops below

the previously set threshold. This point, if it is found, demonstrates that the speech begins with an unvoiced sound and allows the algorithm to return a starting point for the speech, which includes any unvoiced section at the start of the phrase.

6. The above process will be repeated for the end of the speech signal to locate an endpoint for the speech.

A sample result is shown in the figure below.

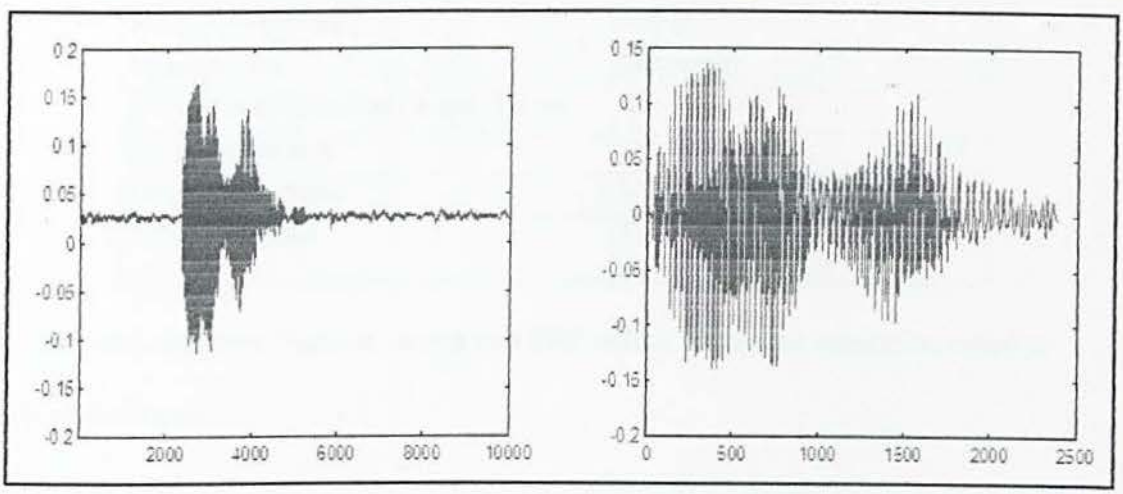


Figure 5.1 Example of Speech Period Extraction

5.3.3 Voice Feature Extraction

MFCC

The Enrollment module was used to call the functions to create the MFCCs. The function implementation of extracting the MFCCs uses a filterbank to extract the power coefficients at specific mel-scale points is *melcepst.m*. Using the filterbank essentially mimics the way the human ear perceives certain frequencies.

The function requires the following parameters: signal, sampling frequency, window type, number of coefficients, number of filters in the filterbank, length of a frame, and the frame increment. Default values are shown in Table 5.1

Table 5.1 Default Mel-Cepstrum Parameters

Parameter	Default Value
Sampling Frequency	16kHz
Window Type	Hamming
Number of Coefficients (in each frame)	12
Filters in filterbank	29
Length of the frame	256
Frame increment	128

Note that the mel-cepstrum function uses a real FFT with a frequency resolution equal to the length of the frame.

5.3.4 Voice Recognition

Vector Quantization

In this project, since little data is available, VQ will be suitable for voice recognition, due to ease of implementation and high accuracy. VQ mapped vectors from a large vector space to a finite number of regions in that space. Each region is called a *cluster* and can be represented by the centroid called codeword. The collection of all codewords consists of the corresponding codebook for a known speaker.

The functions for vector quantization are *vqlbg.m* and *kmeans.m*. Given a set of MFCCs, and the number of desired codeword, *vqlbg.m* outputs a list of codewords. The

function *kmeans.m* is called by *vqlbg.m* to perform the iterations of K-nearest neighbor for the amount of desired codebook size. The distance from a vector to the closest codeword is called distortion. In the recognition phase, an input utterance of an unknown voice is vector-quantized using each trained codebook and the total VQ distortion is computed. The speaker corresponding to the VQ codebook with smallest total distortion is identified.

Figure 5.3 shows, in a flow diagram, the detailed steps of the LBG algorithm used in the coding *vqlbg.m*. "Cluster vectors" is the nearest-neighbor search procedure which assigns each training vector to a cluster associated with the closest codeword. "Find centroids" is the centroid update procedure. "Compute D (distortion)" sums the distances of all training vectors in the nearest-neighbor search so as to determine whether the procedure has converged.

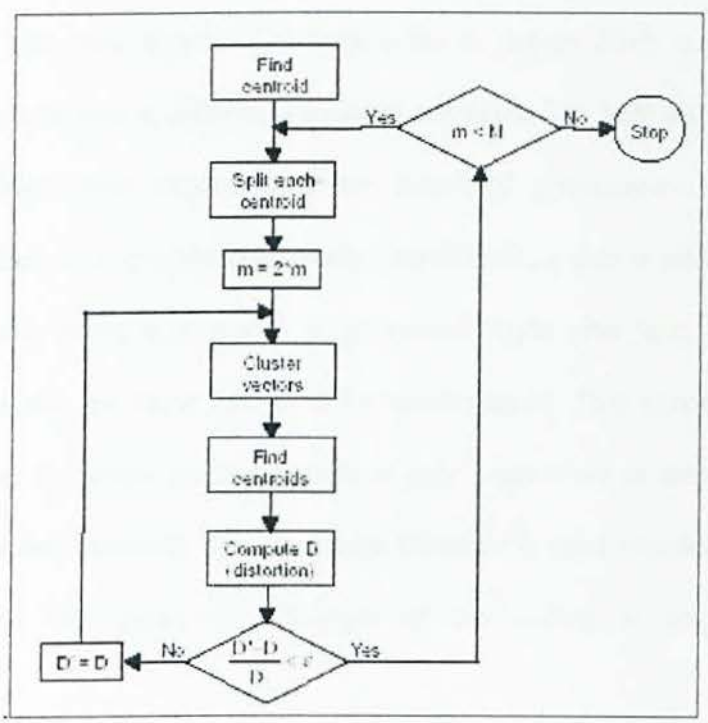


Figure 5.2 Flow-Chart of the VQ-LBG Algorithm

Below is a snippet of the coding

```
nc=size(d,2);
[x,esq,j]=kmeans(d,1);
m=1;
while m<k
  n=min(m,k-m);
  m=m+n;
  e=1e-4*sqrt(esq)*rand(1,nc);
  [x,esq,j]=kmeans(d,m,[x(1:n,:)+e(ones(n,1),:);
   x(1:n,:)-e(ones(n,1),:); x(n+1:m-n,:)]);
end
```

5.3.5 Threshold Creation

A global and static threshold is impossible to define. Each codebook represent different enrollee and thus a different threshold are needed to help aid this situation. A third pseudo-module was implemented for threshold generation. The function for threshold generation is *id_threshold.m.* During enrollment, a user is asked to record once his / her password, where a codebook is generated. Right after that, the same user is asked to record again the same password for confirmation. This second recording is a very effective way to collect another sample of user's password or another codeword in order to calculate the threshold. The Euclidean Distance is used to calculate the average distance of these two codewords. Snippet of the coding is retrieved from the *id_threshold.m.*

```
d = disteu(tstceps', codebook');
distort = sum(min(d,[],2)) / size(d,1);

thresh = (distort*1.2);

idfname = strcat(UserID,'.mat');
%save(idfname,'codebook1','codebook2','codebook3');
save(idfname,'codebook','thresh');

%display(thresh);

end;
```

5.3.6 Decision

The authentication decision is made by calculating the average Euclidean distance between the test vectors and codeword. This is done in the Authentication module, performed by the function *id_test_authen.m* which calls the Euclidean distance calculation function. The average Euclidean distance is then compared to the user's threshold for the pass/fail verdict. Snippet of the coding is shown below.

```
idfname = strcat(UserID,'.mat');        %get userID
load(idfname)                           %load file name
distort = id_test_authen('UserAuthen.wav',codebook);        %return distortion

if distort < thresh
```

```
    set(handles.MsgTxt,'String','Voice Recognition Successful!!!... User Has Been
                        Authentified');
else
    set(handles.MsgTxt,'String','Voice Recognition Failed!!!... Please Try Again');
end;
```

5.4 User Interface Development.

The user interface for the voice authentication system was developed using Matlab 6.5 GUI. The process of creating a more friendly than the actual proposal were somewhat simplified using this software. This system does not require extensive graphical representation as the target of the project is to have a system that can recognize and authenticate user's voice. Nevertheless, a minimal use of color is added to differentiate some functional buttons and thus eliminate a wholly static appearance. The screen shots of the user interface created in the Matlab GUI editor can be seen in Appendix B.

Every function in the interface must be declared correctly. Wrong callback to function might result in system error. Thus, Property Inspector, an important feature of Matlab GUI in interface creation was used to declare every callbacks, tag names, and also styles of the functions. Below are lists to data declared and stored using the property inspector. The bolded text in the top-left side of the table indicates the interface name.

Table 5.2 Main Menu's Property Inspector

Main Menu	Style	Callback	Tag	% Comment
Enrollment	Pushbutton	EnrollBtn	EnrollBtn	Go to enrollment
Authentication	Pushbutton	AuthenBtn	AuthenBtn	Go to authentication
Exit	Pushbutton	ExitBtn	ExitBtn	Exit interface
Static Box	Frame	Text	WelcomeTag	Display static message

Table 5.3 Test Mic's Property Inspector

Test Mic	Style	Callback	Tag	% Comment
Record	Pushbutton	TestRecBtn	TestRecBtn	Record signal
Done	Pushbutton	DoneBtn	DoneBtn	Exit to main menu
Message Box	Text	MsgTxt	MsgTxt	Display messages
Static Box	Text	-	MsgTxt2	Display static message
Signal Box	-	-	Axes1	Display signal

Table 5.4 Enrollment's Property Inspector

Enrollment	Style	Callback	Tag	% Comment
Login ID	Edit	UserIDTxt	UserIDTxt	Get user ID
Record	PushButton	RecBtn1	RecBtn1	Record password
Confirm	PushButton	RecBtn2	RecBtn2	Confirm password
Playback	PushButton	PlayBtn	PlayBtn	Playback password
Back To Main	PushButton	BackBtn	BackBtn	Back to main menu
Exit	PushButton	ExitBtn1	ExitBtn1	Exit interface
Message Box	Text	-	MsgTag	Display message
Signal Box	-	-	Signalaxes	Display signal

Table 5.5 Authentication's Property Inspector

Authentication	Style	Callback	Tag	% Comment
Login ID	Edit	UserIDTag1	UserIDTag1	Get user ID
Record	PushButton	RecordBtn1	RecordBtn1	Record password
Back To Main	PushButton	BackBtn	BackBtn	Back to main menu
Exit	PushButton	ExitBtn	ExitBtn	Exit interface
Message Box	Text	-	MsgTxt	Display message
Signal Box	-	-	Signalaxes1	Display signal

5.5 Chapter 5 Summary

This chapter shows how system was implemented using the software, and technology chosen, as well as the hardware components utilized. The testing done is explained in the next chapter.

CHAPTER 6

-TESTING-

6.1 Introduction to Testing

Testing is an important phase in software development as aside from identifying and rectifying bugs; this phase also determines the quality of the system built. By comparing actual behavior versus expected results, the developer attains a feel of system functionalities as a user. Aside from testing done on the developers part, future users also partake in testing evaluating prototypes of the system, thus helping identify drawbacks and flaws, which requires analysis of the system be revised.

6.2 Stages of Testing

Different stages of testing were used to test the code and functionality of the System Login Authentication. The purpose of testing the code is to ensure system is reliable and stable, whereas the purpose of testing the functionality is test the performance of the system.

During implementation, all codes and programs were compiled on the personal computer at home and at the faculty. There are six different categories of testing, which are: Unit/Component Testing, Integration Testing, System Testing, User Acceptance Testing, Regression Testing and Code Inspection.

6.2.1 *Unit/Component Testing*

Unit testing was done by testing a small piece of code that executes simple tasks like calling another file or displaying another interface. It is non reliant on other units of codes. Unit testing done tested individual functionalities non-dependent on other functionalities. Individual unit interfaces were also tested for usability.

The component testing is the next level of testing performed. Components of a module were also tested to see if each component function met objectives. One example on how this is done is by recording voice inputs and checking whether the files have been generated and named according to the different logged –in ID in the database. The calling function must also generate the same ID codebook whenever the same ID login is detected. Source of input will only be detected if the microphone is used also the output source should only be the speaker and not any other device.

6.2.2 *Integration Testing*

Integration testing is done to ensure that all components performed correctly without problems when integrated. In this case, integration testing was done when both sub-systems of System Login Authentication which are the User Enrollment and User Authentication were integrated. A lot of testing on the calling functions was done to ensure that all callback functions worked smoothly and that each link, function and file could be called and opened successfully. Aside from that, integration testing here also ensured that all modules could access the same database for their respective usage of

codebooks. During this phase, test scripts are generated to see if every adjacent subsystem communicated properly. The System Analysis and Design phase was a big help of reference since the input/output specifications in each module were clearly explained which motivate easy integration. Whenever any callback did not work, the Matlab command prompt was always reliable to refer to as it would display error messages without having to inspect the function codings all over again.

6.2.3 *System Testing*

The goal was to verify that User Enrollment session, and User Authentication session could both work adequately. Both prototypes did not run properly on the first test. This was due to an error in the codings where certain variables were not clearly declared, resulting a confusion to the storage of codebooks. What happens was one of the running loop of codes did not stop and it keeps updating the same codebook instead of creating another. Thanks to the methodology used, the prototyping model, the codings has been revised and enhanced to fulfill system requirements.

The second testing was partly successful. Codebook was generated successfully during the first recording of the User Enrollment module, the threshold was created successfully during the second recording and the results were stored according to the right codebook that bears the right user ID. An average distortion factor was generated properly in the User Authentication prototype. Only the verdict subsystem was disabled for the time being due to uncertainty of the average distortion factor. After a close observation a value was determined and the prototypes were updated.

6.2.4 *User Acceptance Testing*

Future prospective users were invited to preliminarily use the system. Their views were taken and considered for changing certain functionalities, and the views that needed more time to implement will be considered for future enhancement and improvement. The table below is to show the test results.

Table 6.1 User Testing Result

Tester	Test 1	Test 2	Test 3	Test 4
Zati (F)	√	√	√	√
Uniza (F)	X	√	√	√
Majin (M)	√	√	X	X
Nuzula (F)	X	√	√	√
Apai (M)	√	√	X	√
Nie Feng (M)	√	√	√	√
Shamsul (M)	X	√	√	√
Metrius (M)	√	X	√	√

Symbol:

√ Recognized / Authenticated

X Not Recognized / Rejected

M Male

F Female

6.2.5 Regression Testing

Regression testing is done when system is tested once each change of code is performed. This is essential in identifying and reducing bugs in system code. Regression testing was done just to eliminate the few bugs identified.

6.2.6 Code Inspection

Code inspection is important to enhance readability, modifiability and understanding of code by other developers. All members of the project did code inspection where each project member thoroughly checked the code of their respective modules. The outcome of this inspection is based on the developers understanding of programming, thus coding developed might differ from the coding of other programmers.

6.3 Problems Encountered During System Testing

The main problem that was found during the system test is the system sometimes stops after a user has spoken into the microphone. An error message was displayed in the Matlab command line, which is '*Matrix dimension does not match*' or '*Index exceeds matrix dimension*'. This, however, is not a fatal error and the user can still click the record button again to re-record their voices until the correct message is displayed. The reason for this problem is probably due to faulty microphone or when the sound recorded exceeds the maximum decibel value of 40 and the signal is too convoluted. Convoluted signal may arise if the user breathes into the microphone while recording his/her

password and also if too many plosive sound is used for his/her password. Currently the problem is not yet solved.

Another occurring problem is the message that represents the verdict of authentication – whether user's voice is recognized or not, sometimes failed to appear on the message box. Other functions still run for example the signal box was displaying every recorded signal but the message box appeared empty instead of displaying texts regarding to system actions at the moment. It was also greatly noted that this problem arises not all the time but more often when the system is run on a lower performance computers. It could be due to the slow system reading of parameters but with a little bit of patience, re-trials from user re-buffered the computer and normally were worth to wait. A conclusion was made – a performance test must be done in order to seek for excellent system executions.

6.4 Performance Test

The goal of this test phase was to find optimal values for certain parameters in the system, the accuracy of the system, speed of execution and disk space usage.

6.4.1 Optimization Tests

The parameters under test included codebook size, number of MFCCs per acoustic vector, and threshold generation scaling factor. Test scripts were written to vary each parameter and observe the average distortion generated each time.

Optimality for the threshold generation scaling factor was determined to minimize false acceptances and false rejections. Tests showed that if the scaling factor was too high

– more than the average 16000Hz or 22050Hz for a 2 second sampling length, there were significantly more false acceptances. And if the scaling factor was too low – lower than 8000Hz, there were significantly more false rejections. This is obvious since the threshold is in the sensitivity level. After a careful testing, a scaling factor of 1.2 was determined to be the optimal value. Optimality for the other two was determined based on a speed vs. accuracy trade-off. The optimal values are summarized in Table 7.1.

Table 6.2 Optimal Parameter Values

Parameter	Optimal Value
No. of MFCCs per acoustic vector	12
No. of Codewords	64
Threshold Generation Scaling Factor	1.2

6.4.2 Execution Time Test

The execution speed of the User Enrollment, Threshold Generation, and User Authentication modules were analyzed. It was found that the time needed to perform the codebook creation in the Enrollment module took the most time compared to the others while the threshold creation took the least time.

6.4.3 Disk Space Usage

The codebook size for a 64-codeword codebook and 12 MFCCs required about 7 kilobytes of disk space. The threshold value for each user required 9 bytes.

Three .wav files are created and reused for the system, which are User1st.wav, User2nd.wav and UserAuthen.wav. Each wave files uses up about 63 kilobytes of space.

6.5 Chapter 6 Summary

The outcome of the testing phase uncovered minor drawbacks and errors, which could be fixed by the developers, none of which required extensive change that made it necessary to change system requirements. It can be concluded that the testing phase was performed carefully to ensure system developed met the objectives and requirements proposed. Future enhancement is proposed for better implementations of system. Several suggestions and recommendations will be carried out in the near future to cater to extra requirements gathered from views by prospective users during the user acceptance testing. This is so because due to project deadline, these few requirements cannot be met just yet.

CHAPTER 7

-SYSTEM EVALUATION-

7.1 Introduction to System Evaluation

The evaluation of this system is concluded by views from the system developers, and also from future prospective users. From these evaluations, the current advantages and disadvantages of developed system has been identified and discussed in the subsections to follow. Future enhancements are proposed to further improvise the current setbacks detected of the system for future use and implementations.

7.2 System's Discussion – Strength & Weakness

7.2.1 Target Requirements

From the developers' point of view, this system has achieved its target requirements as providing a means of authentication to system login users by recognizing their voice. From survey done during the analysis phase, there were not many Malaysian born systems that explore this new exciting field of biometric technology of voice recognition. Thus in the developers eye, one of the strength of System Login Authentication Using Voice Recognition is it has achieved its objectives of exploring a new domain of artificial intelligence discovery and computer access security.

7.2.2 Enrollment Module

This module was developed with the goal of using a chosen method to be used in recognizing a user's voice. Creation of codebook using the same ID as the enrolling user and generations of .mat and .wav files for the same user in the database has been successful. Since wave files (User1st.wav and User2nd.wav) have been declared globally, every new enrollment will update and replace the old wave files. There is no necessity for

the system to keep stored two extra files for each enrollee since the codebook generated is enough to keep distinct information of every user, plus this will boost a bigger capacity for storage purpose. Nevertheless, these updated wave files are important to the playback function. Without them, the playback function cannot output an analog signal since most audio device detect wave files as the minimal standard for audio plays. This playback functions as a reminder of the password just enrolled because some user might want to confirm it is their voice that has been recorded not something or someone else's plus it help improve user's memory over time due to human cognitive limitations.

Creation of threshold for every user is generated automatically by multiplying the average distortion (the average euclidean distance between the test vectors and the codebook vectors) by 1.2. This is a fixed scaling value. There was consideration to expand it to three extra optional values – low, normal and high, and let the user decide his/her owns but after reviewing the analysis of system requirements, the idea was rejected due to the fact that a security system should provide the same level of security for every single user. It is also important for user to rephrase the password with fair accuracy as their first recordings in order to not have a big ratio in the threshold less imposter will easily authenticate themselves.

7.2.3 Authentication Module

The authentication module is where the pass/fail verdict is given. It is a resulting comparison of average distortion values of feature vectors and the threshold. A message displaying successful recognition will appear for authentic users and vice versa for none authentic trials. For rejected authentications, users can only submit to repeat the authentication process until it works or choose to re-enroll themselves again.

Based on the results obtained during testing, it was observed that the False Rejection Rate (FRR) where the authentic speaker is rejected is quite high. The False Acceptance Rate (FAR) where an impostor is verified successfully, however, is very low. Referring to past researches, it was learned that the cause of the high rejection rate was probably due to the length of the user's spoken password. A longer utterance will normally yield to a less FRR.

7.2.4 Microphone Issues

A very sensitive microphone often results in the problem that was discussed in section 6.2.7. If it occurs, re-clicking the record button and re-record the voice seemed to be the only way to get to the desired results. Seldom there requirements to change to a new microphone but nevertheless it does stand as an option. Try to avoid breathing into the microphone because this will convolute the speech signal. Currently, there is no other way around this problem.

7.3 Future Enhancements

This section achieved its ideas from the views and opinions gathered from the drawbacks section. Hence developers have proposed some enhancement that can be done in the future. For further explanation and detail information, it is recommended to use the list of reference and bibliography.

7.3.1 Feature Extraction/Representation

- RASTA-PLP

Another popular speech feature representation is known as RASTA-PLP, an acronym for Relative Spectral Transform - Perceptual Linear Prediction. PLP was originally proposed by Hynek Hermansky as a way of warping spectra to minimize the differences between speakers while preserving the important speech information. RASTA is a separate technique that applies a band-pass filter to the energy in each frequency sub band in order to smooth over short-term noise variations and to remove any constant offset resulting from static spectral coloration in the speech channel e.g. from a telephone line. [8]

7.3.2 Recognition Alternatives

- Dynamic Time Warping

Deals with inconsistencies in the sampling rate of speech by stretching or compressing parts of the signal in the time domain. Dynamic time warping addresses timing misalignments between signals. One of the difficulties in speech recognition is that although different recordings of the same words may include more or less the same sounds in the same order, the precise timing - the durations of each sub word within the word - will not match. As a result, efforts to recognize words by matching them to templates will give inaccurate results if there is no temporal alignment. Although it has been largely superseded by hidden Markov models, early speech recognizers used a dynamic-programming technique called Dynamic Time Warping (DTW) to accommodate differences in timing between sample words and templates. The basic principle is to allow a range of 'steps' in the space of (time frames in sample, time frames in template) and to find the path through that space that maximizes the local match between the aligned time frames, subject to the constraints implicit in the allowable steps.

The total 'similarity cost' found by this algorithm is a good indication of how well the sample and template match, which can be used to choose the best-matching template.

- Hidden Markov Models

For speech signals, a left-right HMM model is found to be more useful. A left right model has the property that as time increases, the state index increases (or stays the same), that is the system states proceed from left to right. Since the properties of a speech signal change over time in a successive manner, this model is very well suited for word recognition. 6-state left-right HMM is used for modeling 2-phoneme password.

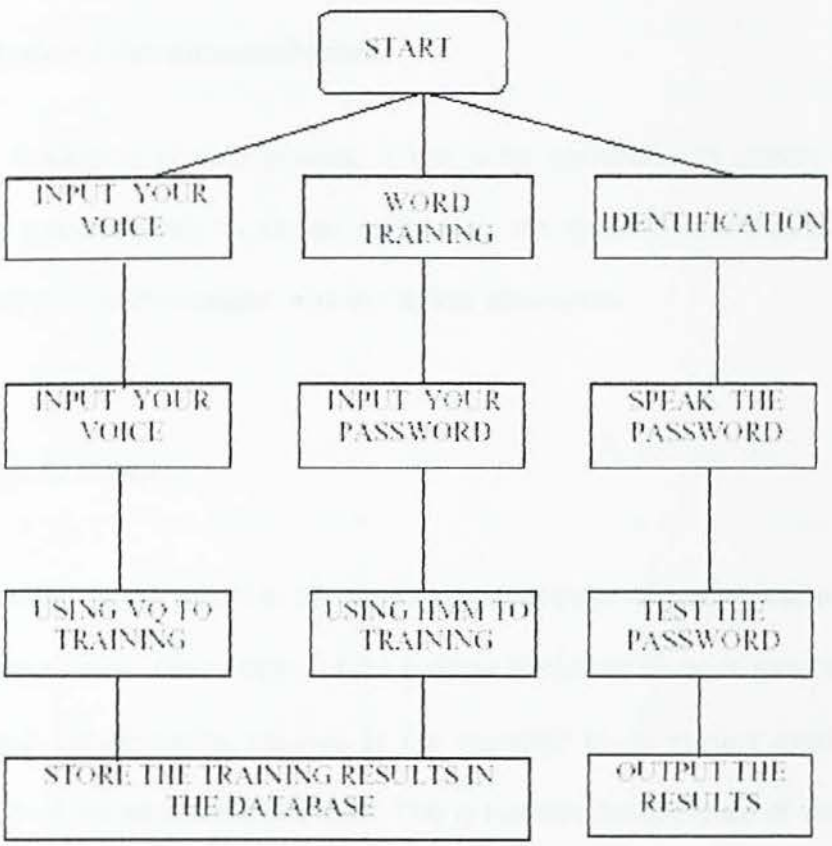


Figure 7.1 The Hidden Markov Model

- Power method

This method uses the power, or energy, properties of the words for recognition. It is easily implemented since it only requires power computations of the tested words. We will use the fact that words are pronounced different and therefore that the power is distributed different 'i' time. By training a model with several realizations of our words and average together, this will give us a specific pattern for each word that we can compare with the word we are searching for.

7.4 Suggestion / Recommendations

Due to the limitation of time project, it was a bit unrealistic to pursue a few more adventurous enhancements for certain modules of the system. Nevertheless, all ideas is given a closer look and discussed with my fellow developers.

7.4.1 Threshold Training

Currently, the threshold is set based on the average distortion calculated by one authentication session. The variance of the average Euclidean distance generated from the threshold and authentication sessions is not expected to be perfect even though the recognition is of the same authentic user. This is because the utterance of words in terms of pitch and resonance of any human over time is never the same from trials to trials. Hence it is suggested that every time a user use the authentication module, the recorded voice is looped and trained with the existing codebook in the database to get a new

updated version of codebook and produce an equal updated version of threshold. The goal is to make sure that a valid user's password that has variation in terms of, again, pitch and resonance is always recognized. Then, it can be well said that the system can not only recognize, but also learn the voice characteristics of an utterance thus authenticate a person.

7.4.2 Code Vectors Weighting

If the codebooks for several users were compared, it can be seen that certain code words are generated relatively close to each other among the different users. Accuracy could be improved by creating some sort of weighting for each code word in each codebook. This way, code words that occur frequently would be weight less in the average distortion calculations.

7.4.3 Signal Normalization

In the Mel-Frequency transform, the output of the filterbanks depends on the power of the signal. This implies that a loud utterance is seen differently than low utterance of the same word. Signal normalization using certain specified method might perhaps reduce this effect.

7.5 Chapter 7 Summary

The views and opinions of this section is essential in listing out the pros and cons of the developed system and as positive encouragement to the developers for the betterment of their login authentication system activities and also to give ideas to further boost the positive impact of the system.

CHAPTER 8

- CONCLUSION-

8.1 System Conclusion

Overall, the project can be considered a success with the basic requirements being satisfied. The finished product could enroll users, verify their voiceprint, and provided a gui interface for users to do so. Matlab 6 is required to run the full program.

Performance of the system was fair with a false rejection rate of 25% and a false acceptance rate of 9.6%. False rejection errors are the result of the system not being able to overlook the small changes in a person's voice recording, for example, the emphasis that the put onto syllables or the changing tone of their voice. False acceptance errors occur when the imposter's voice has similar frequency characteristics to the true user. Adding threshold generation using multiple recordings and weighting of the code vectors would improve both false acceptance and rejection ratios while codebook adaptation and signal normalization would improve false rejection ratios.

8.2 Project Summary

This project is a requirement of the course taken at the Faculty of Computer Science & Information Technology, University Malaya, and is an important requirement to be fulfilled in order to graduate for the course taken.

This project was developed in the time span of two semesters, where the first semester concentrated fully on the project proposal, reviewing available literature,

eliciting user requirements and designing the system. The second semester consisted wholly on the system implementation, testing and evaluation phase. The whole project was conducted using the traditional software development life cycle phases.

In the whole two semesters of project development and implementation, the developer has learnt much and gains valuable experience regarding software development, even though the system developed will not be considered a big project in a real world-working environment. Nevertheless being wholly involved and responsible for all phases of system development has proved invaluable and is essential experience. Without really realizing, the past three years of education were the main guidance needed in developing and implementation of this project. It also cannot be denied that being constantly aware of technological advancements, being inquisitive and also seeking guidance from those more experienced are also important in the learning phase of project development and implementation.

In the development process, it's unavoidable to experience some setbacks and problems. In this case, the developer experienced a major system crash of her personal pc at the faculty, due to virus infection. Thus it is so very important to execute and keep updated backups frequently.

The initial phase of learning and having complete understanding of new technology proposed for use of the developed system was also initially frustrating, as there was so much to learn and not enough time. But with patience, guidance and encouragement, a system developer can overcome this. The developer also recognized

the importance of constant and good communication with partner developers as in a group project, such as this communication is essential.

As a whole it can only be repeated and concluded again, that even though system development is not a piece of cake, at the end of the day, it provides satisfaction from seeing the system developed finally implemented and gathering experience and knowledge that has proved invaluable.

APPENDICE

Appendix A

Terminology

Several terms used throughout this document are clarified here:

1. Authenticate:

Establish or prove as:

a). Conforming to a fact and therefore worthy of trust, reliance, or belief.

b). Having an undisputed origin.

c). To establish the validity of a claimed identity

2. Biometric:

A measurable, unique physiological characteristic or behavioral trait used to recognize the identity, or verify the claimed identity, of an enrollee(i.e a subject or user).

3. Biometric Image:

The raw (unprocessed) output of the biometric scanner. The biometric image is generally fed into a biometric extraction algorithm.

4.Biometric Template:

A data set representing the biometric measurement of an enrollee which is maintained on file and used by a biometric verification device for comparison against subsequently submitted biometric samples. The biometric template is generally the output of biometric extraction routine(s), such as a Minutia Extraction Algorithm.

5. Certify:

To confirm formally as true, accurate, or genuine.

6. Certificate:

A document testifying to accuracy or truth.

7. Enrollment:

The process of collecting biometric samples from a person and the preparation and storage as a template of that person's identity

8. False Acceptance:

When a biometric transaction results in the acceptance of an imposter (also known as a False Match or Type II Error).

9.False Rejection:

When a biometric transaction results in the failure to recognize the identity, or verify the claimed identity, of an enrollee (Also known as a False Non Match or Type I Error).

10. Identify:

a)To establish that the collective aspects of the characteristics by which a thing is distinctly recognizable or known.

b)To consider similar or identical: EQUATE.

11. Imposter:

A person who submits a biometric sample in either an intentional or inadvertent attempt to pass him/herself off as another person who is an enrollee.

12. Verify:

To prove the truth of by presenting evidence or testimony: SUBSTANTIATE.

13. Privilege:

A special grant, immunity, right or benefit granted to an individual, class or caste.

14. Response time:

The time period required by a biometric verification device to complete a biometric transaction.

15. Security Policy:

The set of laws, rules, and practices that regulate how an organization manages, protects, and distributes sensitive information

16. Subject:

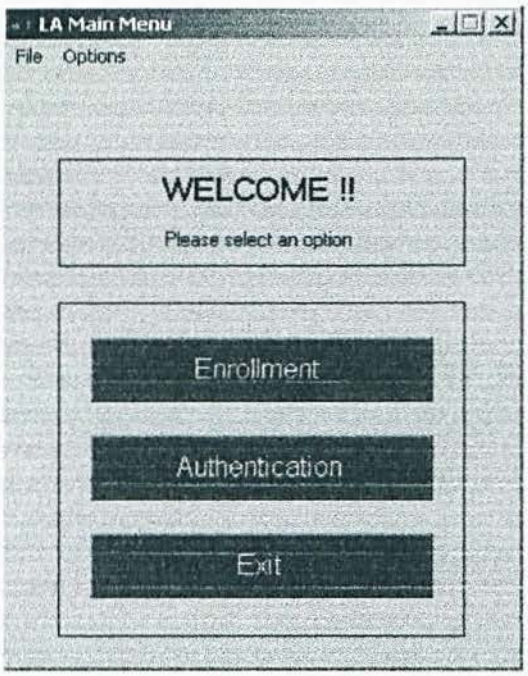
An active entity, generally in the form of a person, process, or device that causes information to flow among objects or changes the system state.

17. Zero Knowledge: a protocol used to establish the identity of a user by proving knowledge of a secret but without revealing that secret. An example is a challenge-response protocol.

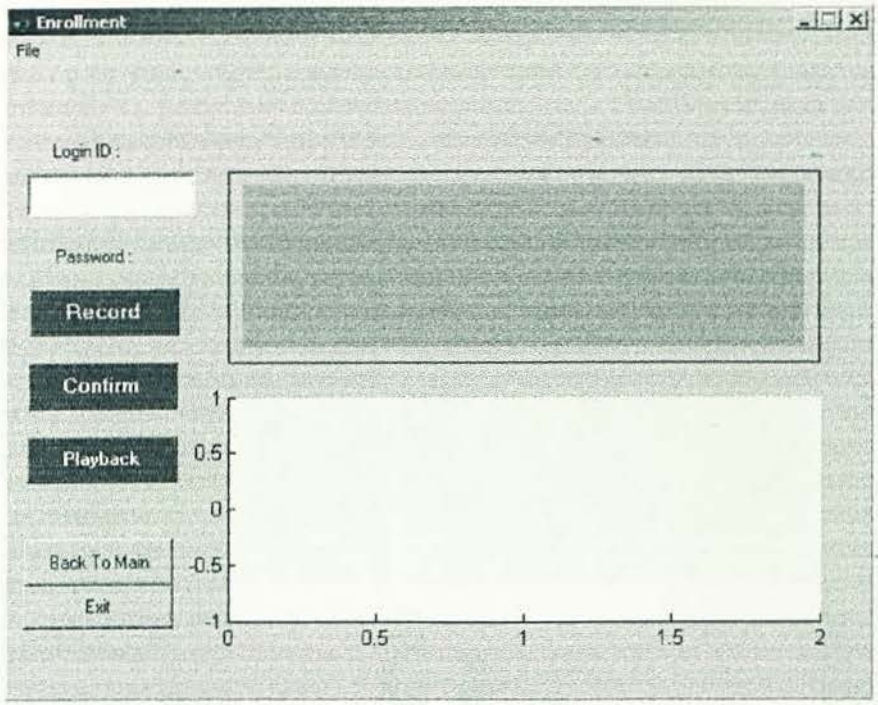
Appendix B

GUI and User Manual

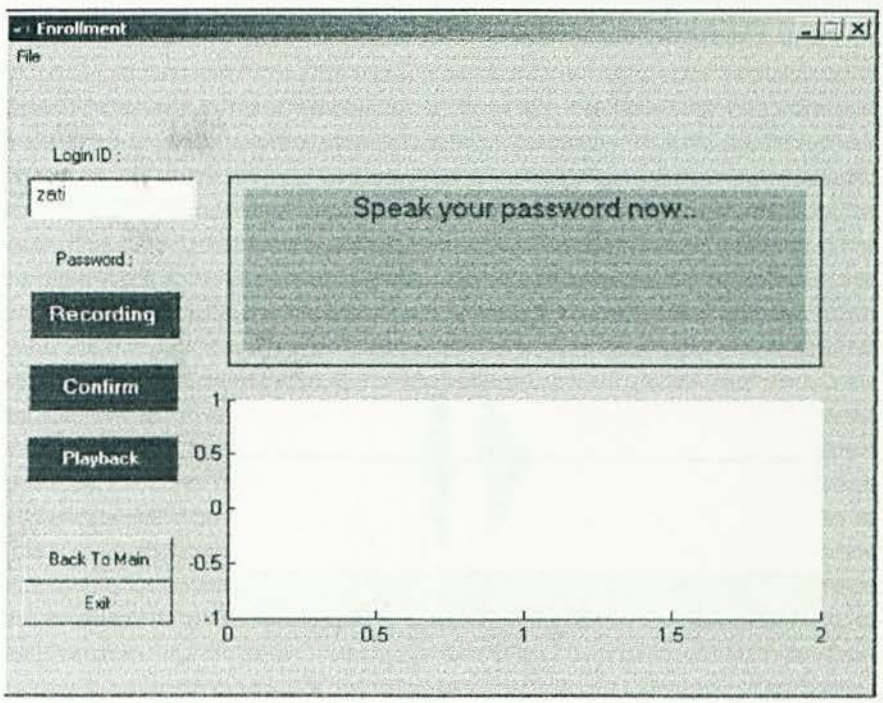
1. Open Matlab and the file MainMenu.fig in the workspace domain. The main menu of the system will appear.



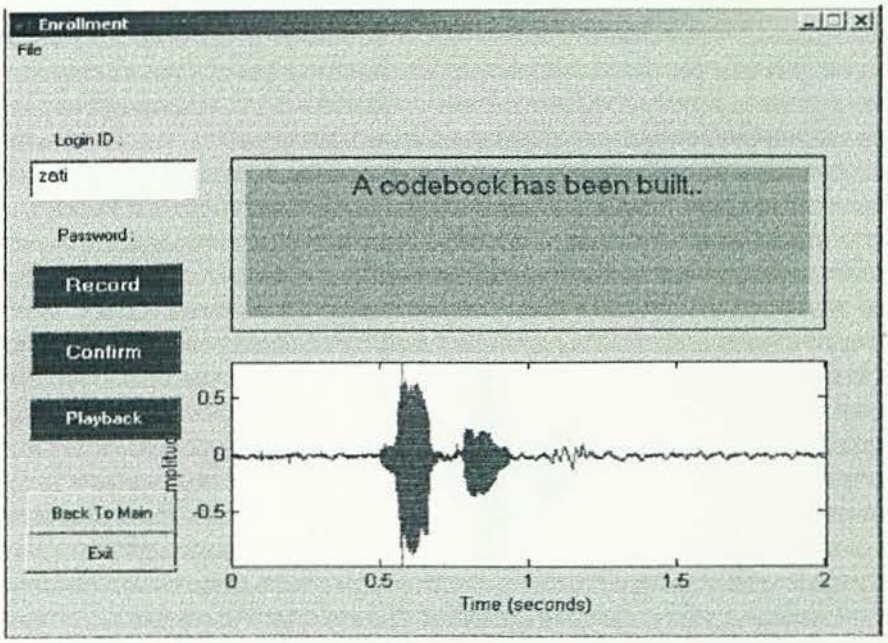
2. Choose which options from the two available module.
   2.1 Enrollment Module
      a. The enrollment screen will be brought up.



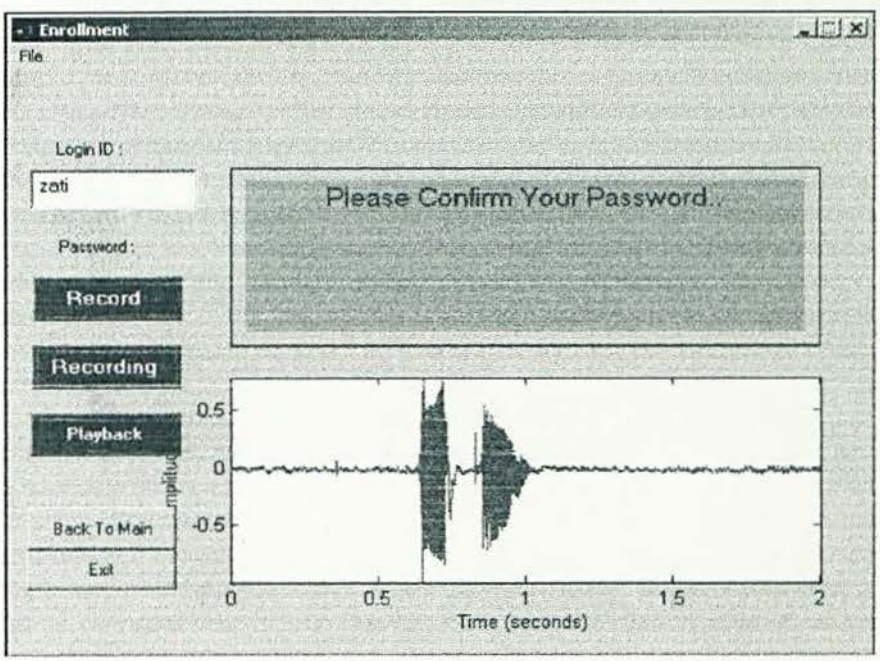
b. For enrollment of new users, enter an ID and then click the Record button. Wait until the message "Speak your password now.." appears before speaking into the microphone. *Note: The duration of recording is2seconds.



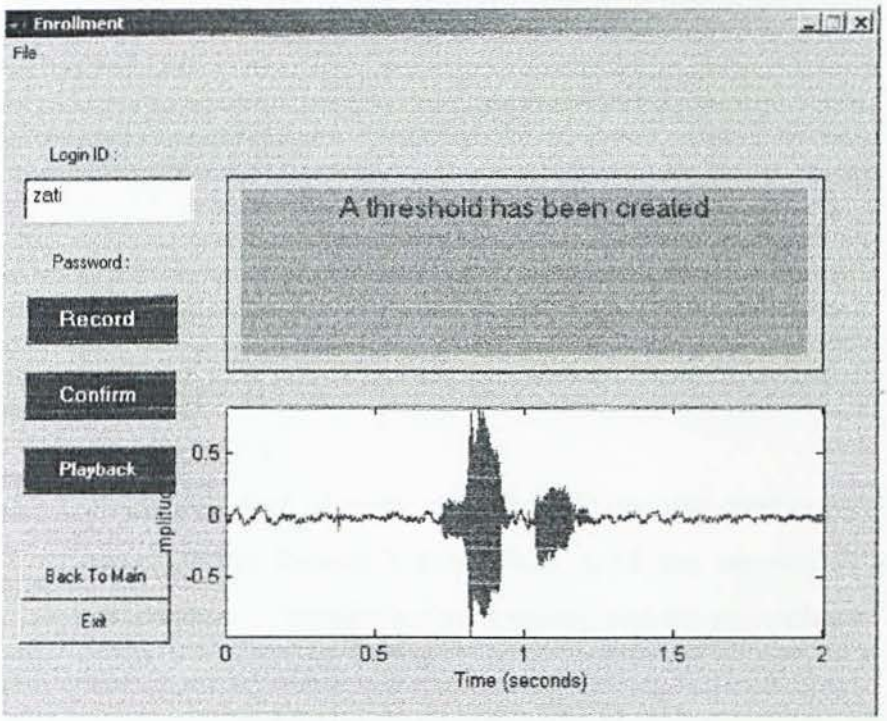
c. The system will perform the necessary operations (It may take awhile) and then the signal waveform will be plotted and the message "A codebook has been built.. " will be displayed.



d. Click the Confirm button and the wait until the message "Please Confirm Your Password .." appears before speaking into the microphone.



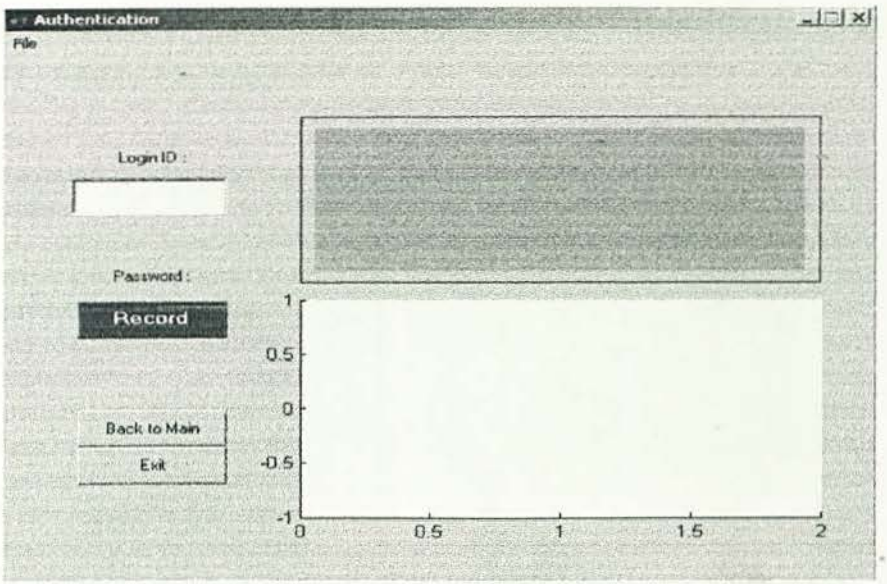
e. Wait until the signal waveform and the message "A threshold has been created" appears.



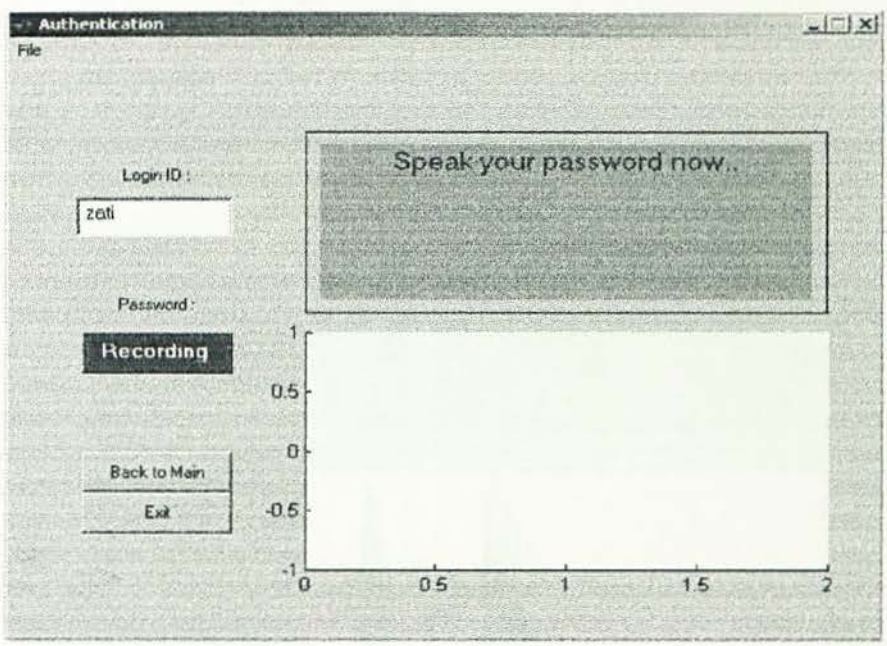
f. Enrollment procedure is completed.
You may use the Playback button to listen to what you've recorded or you may choose other options.

2.2 Authentication Module

a. The authentication screen will be brought up.

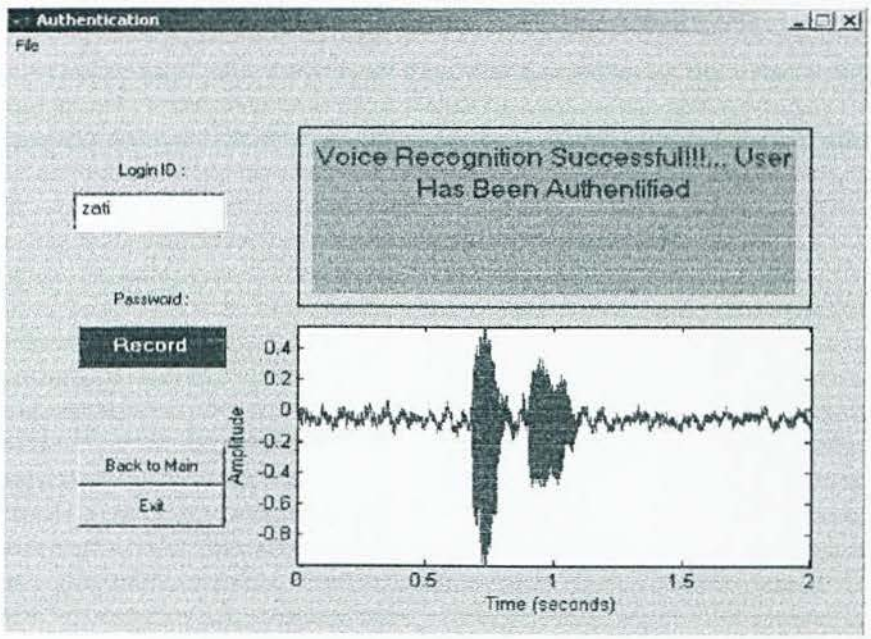


b. For authentication of users, enter their ID that was used in enrollment and then click the Record button. Wait until the message "Speak your password now.." appears before speaking into the microphone.

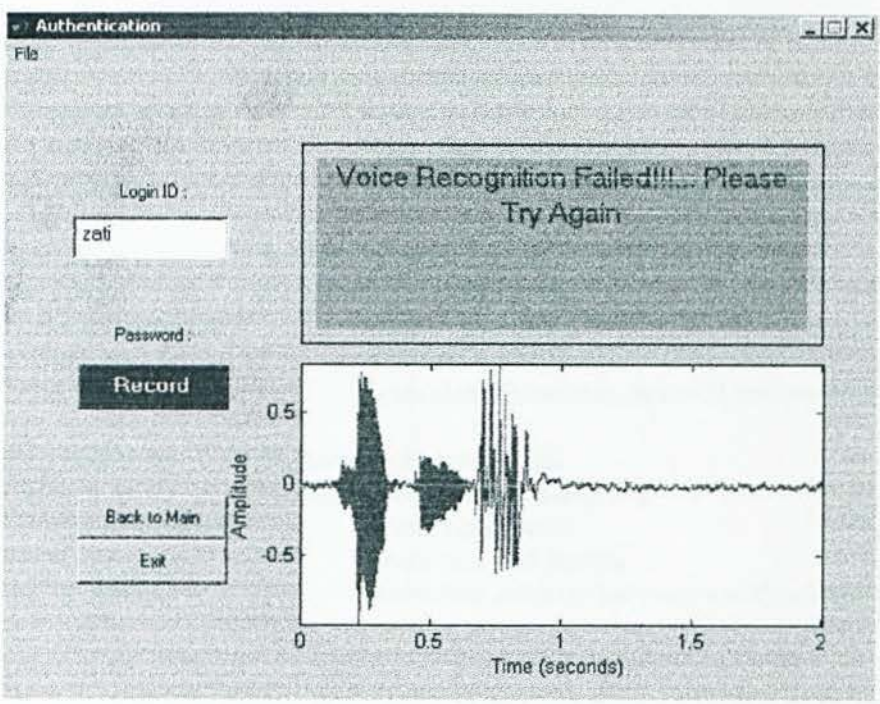


c. The system will perform the necessary authentication operation. If the authentication is successful the signal waveform will be plotted and the

message "Voice Recognition Successful!!!..User Has Been Authentified" will be displayed.



Else the message "Voice Recognition Failed!!!... Please Try Again..." will appear.



d. The authentication is finished.

Appendix C

Functions and Call Structures

The following is the hierarchy of the main files that are called in the system. Some files are listed twice because they are called from different places. All data files have the extension *.mat*.

Graphical User Interface

MainMenu.fig, mainmenu.m

Enrollment.fig, enrollment.m

Authentication.fig, authentication.m

TestMic.fig, testmic.m

Core Functions

```
mainmenu.m              %start calling function

  enrollment.m          %return userID, codebook and threshold
      userID                    %get userID
      init_sound.m              %get analog input / signal

      id_train_speaker.m    %return the codebook
      test_train            %compute training using VQ
          locatespeech.m            %detect end-point
          melcepst.m                %calculate signal feature
          vqlbg.m                   %calculate Euclidean distance of codebook

      id_treshold.m         %return threshold value
      test_threshold        %compute threshold using VQ
          locatespeech.m            %detect end-point
          melcepst.m                %extract signal feature
          disteu.m                  %calculate distance codeword1,codeword2)

      id_test_authen.m      %return verdict pass / fail
      test_authen.m         %compute comparison using VQ
          locatespeech.m            %detect end-point
          melcepst.m                %extract signal feature
          disteu.m                  %calculate distance between codebook and threshold
```

Subfunctions

1. locatespeech.m %end-point detection
    removedc.m %remove dc offset / silence
    avgmag.m %get average magnitude of signal
    zerocrossing.m %get zero-crossing rate of signal

2. melcepst.m %return extracted signal feature
    enframe.m %split signal into overlapping frames
    rfft.m %compress data into specified length
    melfb.m %determine matrix for mel-spaced filterbank
    Melbankm.m %determine matrix for mel-spaced filterbank
    rdct.m %transform data to truncated length

3. vqlbg.m %classify/mapped vectors into regions
    kmeans.m %find nearest-neighbour
        disteusq.m %find distortion between neighbouring data

4. Timedata.m %plots time data at location specified by user
    Waveplot.m %plot axis (x and y)
    Plot scatter.m %draw blocks

REFERENCE

Reference

[1] Rabiner L. and Juang B.H, Fundamental of speech reeognition, Prentice Hall, 1993.

[2] M. D. Skowronski and J. G. Harris, "Human factor cepstral coefficients," *IEEE Trans. Speech and AudioProcessing*, Submitted July 2002.

[3] Fredric M. Ham and Ivica Kostanic, "*Principles of Neurocomputing for Science and Engineering*,"McGraw-Hill, 2000

[4] David Pitts, Bill Ball, et al, " *Red Hat Linux 6*," SAMS, 1999

[5] Alan V. Oppenheim, Alan S. Willsky, "*Signals And Systems*," Prentice Hall, 1983

[6] Tom M.Mitchell, " *Machine Learning*", McGraw-Hill, 1997

[7] Timothy J.Ross , " *Fuzzy Logic With Engineering Applications*," McGraw-Hill, 1995

[8] H. Hermansky, "Perceptual linear predictive (PLP) analysis of speech", J. Acoust. Soc. Am., vol. 87, no. 4, pp. 1738-1752, Apr. 1990.

Title : Terminologies and Definitions

Retrieve : March 7th 2003

From : The World Wide Web

1) http://www.sm-e-commerce.org/glossary.html
2) http://www.grpmax.com/topics/aiexplained1.htm
3) http://nene.essortment.com/informationabou_rgeu.htm

Title : Biometrics

Retrieve : March 13th 2003

From : The World Wide Web

1) http://www.asee.org/nrl/labs/nrl-5500.cfm
2) http://www.admsyst.com/comparison_bio.htm
3) http://service.boulder.ibm.com/software/pervasive/info/EmbeddedVVMultiplatformEd_final.pdf

Title : Voiceware Products

Retrieve : March 28th 2003

From : The World Wide Web

1) http://www.halfbakery.com/idea/Couch_20Potato_20Voice_20Recognition
2) http://www.etechkorea.info/articles/20020307001.php
3) http://www2.whidbey.com/lighthook/cyber5.htm

Title : Feature Analysis

Retrieve : March 27th 2003

From : The World Wide Web

1) http://neural.cs.nthu.edu.tw/jang/research/web/speaker/
2) http://www.hpl.hp.com/techreports/2002/HPL-2002-43.pdf
3) http://www.computerworld.com/computerworld/records/whitepapers/circle_wp.pdf
4) http://www.bbc.co.uk/rd/pubs/whp/whp-pdf-files/WHP015.pdf
5) http://www.voicegenie.com/pdf/PowerToThePeople,11.26.01.pdf

6) http://www.cgchannel.com/news/showfeature.jsp?newsid=1236

7) http://www.admsyst.com/voice/voice.htm

8) http://www.voiceautomated.com

9) http://www.voicepower.co.uk

10) http://hometown.aol.com/tomu1/vrt.html

Title : Authentication / Verification Technologies

Retrieve : March 31st 2003

From : The World Wide Web

1) http://www.3dspacecadet.com/

2) http://www.graymatter.co.nz/static/

3) http://www.cse.iitk.ac.in/users/langtech/strans2002/mohanti.pdf

4) http://www.grad.cmu.ac.th/abstract/2000/sci/abstract/sci08007.html

Title : AI Solutions in Voice Recognitions

Retrieve : April 2nd 2003

From : The World Wide Web

1) http://home.ipoline.com/~timlin/neural/NeuralNetwork.doc

2) http://www.generation5.org/aisolutions/bio00.shtml

3) http://www.seattlerobotics.org/encoder/nov98/neural.html

4) http://www.tradetrek.com/education/ai/ai_stock_trading03.asp

Title : Voice Recognition and Authentication

Retrieve : April 20th 2003

From : The World Wide Web

1) http://webtools.cityu.edu.hk/news/newslett/voicerecognition.htm
2) http://www.ucalgary.ca/~jha/voice2.html
3) www.lce.hut.fi/~vmakine2/mscc.pdf
4) www.ee.columbia.edu/~dpwe/resources/matlab/rastamat/

Title : Digital Signal Processing

Retrieve : April 23rd 2003

From : The World Wide Web

1) http://www.owlnet.rice.edu/~elec532/PROJECTS98/ speech/cepstrum/cepstrum.html
2) www.owlnet.rice.edu/~elec301/Projects01/ speaker_id/trainarch.html
3) www.hasanayaz.com/dsp/473pr.pdf
4) http://www.spsc.inw.tugraz.at/courses/speechlab/ speech_analysis.pdf
5) www.cnel.ufl.edu/~markskow/papers/iscas03.pdf
6) http://research.compaq.com/downloads.html.
7) http://spib.rice.edu/spib/select noise.html.