SUDAN UNIVERSITY OF SCIENCE &TECHNOLOGY

FACULTY OF COMPUTER SCIENCE & INFORMATION TECHNOLOGY

DEPARTMENT OF SOFTWARE ENGINEERING

UAV CONTROL AND SIMULATION USING VOICE RECOGNITION

تحكم ومحاكاة طائرة بدون طيار بإستخدام تقنية التعرف علي الصوت

Oct 2017

A research submitted in Partial Fulfillment of the Requirement for the Degree of Bachelor in computer science

بسم الله الرحمن الرحيم

SUDAN UNIVERSITY OF SCIENCE &TECHNOLOGY

FACULTY OF COMPUTER SCIENCE & INFORMATION TECHNOLOGY

DEPARTMENT OF SOFTWARE ENGINEERING

UAV CONTROL AND SIMULATION USING VOICE RECOGNITION

تحكم ومحاكاة طائرة بدون طيار بإستخدام تقنية التعرف علي الصوت

Oct 2017

A research Submitted in Partial Fulfillment of the Requirement for the Degree of Bachelor in computer science

PREPARED BY:

Asma Eshag Ahmed Saied

Baraa Mohammed Modthther ALSheik

Salma Abdelgader Abdelhameed Mohammed

Hiba Hassan Ali ALnour

الآية

قال تعالي:

بسم الله الرحمن الرحيم

اقرأ بسم ربك الذي خلق (*) خلق الإنسان من علق (*) اقرأ

وربك الأكرم (*) الذي علم بالقلم (*) علم الإنسان

مالم يعلم(*)

الآهداء

إلى من تتسابق الكلمات لتخرج معبرة ع مكنون ذاتها من علمتني معنى الحياة ، وعانت الصعاب لأصل إلى ما أنا عليه الآن ، إلى ملاذي الآمن الذي ألجأ إليه عندما تكسوني الهموم و اسبح في بحر حنانها ليخفف من آلامي، إلى نبضي وذاتي (أمي الحنونه) .

إلى من تجرع الكأس فارغاً ليسقيني قطرة حبٍ الي من حصد الأشواك عن دربي ليمهد لي طريق العلم ، إلى من أحمل أسمه بكل إفتخار ، إلى قدوتي وضياء طريقي ،إلى سندي (أبي الغالي) .

إلى من يضيئون لي الطريق ويساندونني ويتنازلون عن حقوقهم لإرضائي والعيش في هناء ، أحبكم حباً إذا مر على ارضٍ قاحلة لتفجرت منها ينابيع المحبة (إخوتي الاعزاء).

الآن تفتح الاشرعة وترفع المرساة لتنطلق السفينة في عرض بحر واسع مظلم هو بحر الحياة وفي هذه الظلمة لا يضئ إلا قنديل الذكريات ذكريات الاخوة البعيدة إلى الارواح التي سكنت روحي الي الذين احببتهم واحبوني ، إلى رفقاء دربي ، إلى من يمسكون بيدي جيداً ويخبرونني أنني أستطيع الوصول ، إلى من يمسحون الحزن عن قلبي ويرسمون إبتسامات الفرح على روحي ، إلى نصفي الآخر (اصدقائي رفقاء دربي).

إلي كل من أشعل شمعة في طريق مسيرنا وعلمنا أن الإجتهاد طريق النجاح ، إلى من يحترقون ليضئوا لنا عتمة الحياة ، إلى من علمونا أن الفشل بداية رحلة النجاح ، وأن السقوط لا يعني النهاية ،إلى من علمونا العطاء (اساتذتي الكرام).

في مثل هذه اللحظات يتوقف اليراع ليفكر قبل أن يخط الحروف ليجمعها في كلمات ..... تتبعثر الأحرف وعبثاً أن نحاول جمعها في سطور ،سطوراً كثيرة تمر في الخيال ولايبقى لنا في نهاية المطاف الا قليلاً من الذكريات وصور تجمعنا برفاق كانو الى جانبنا وشاركونا جميع اللحظات ، إلى قناديل الفرح ، إلى كل من كان حاضراً وغائباً ، إلى كل من جمعتنا بهم محطات الحياة

لكم جميعا .....نهدي ثمرة جهدنا هذه .

شكر وعرفان

إلهي لايطيب الليل إلا بشكرك ولا يطيب النهار إلا بطاعتك ..ولاتطيب اللحظات إلا بذكرك .. ولاتطيب الأخرة ألا بعفوك ..ولاتطيب الجنة ألا برؤيتك لك الحمد والشكر دائماً وأبداً

الشكر الجزيل إلى الصرح الشامخ ، منهل العلم ، جامعة السودان للعلوم والتكنولوجيا .

إلى الأم الحنون ، المربية الفاضلة ، إلى من كانت معنا خطوة بخطوة

الاستاذة الجليلة د. هويدا عبدالقادر،

إلى من سهر الليالى الطوال مرشداً ودليلا ، إلى من حمل المسؤلية بكل صبر وثبات

أ.شريف

إلى جميع الأساتذة الأجلاء

ورفقاء العلم الأعزاء

المستخلص

يواجه مشغل الطائرة من غير طيار صعوبة في توجية عدت طائرات في اني واحد بالإضافة الي أنه يتطلب الدراية الكاملة بنوع المتحكم المستخدم في الطائرة وديناميكية الطائرة .

الغرض من هذه الرسالة التحكم في الطائرة من خلال التعرف على الصوت لتسهيل هذه المهمة على المشغل والهدف الاساسي من هذه الرسالة تطبيق الشبكات العصبية لأداء هذه المهمة وتصميم ومحاكاة نظام التحكم للطائرة ، تم محاكات التصميم على برنامج البروتوز والماتلاب بإستخدام المتحكم الدقيق ات ميقا 16

تعرفة الشبكة بنسبة 98.3% على الاصوات المدخلة وتم الحصول على نتائج جيده لحركة الطائرة بناء على نتائج الشبكة .

Abstract

Guiding multiple UAVs through keyboard and remote control is very difficult to perform by one operator and it requires the operator to be professional of the specific controller and UAV dynamic behavior. This thesis aims to control UAV using the speech recognition to facilitate this mission. The main objective of this thesis is to apply neural network to perform this task and to design and simulate control system of UAV. The design is simulated in PROTUES and MATLAB/SIMULINK using Atmega 16 microcontroller. The network was well recognize and its accuracy is 98.3%. Good results and desired motion of UAV accordingly were obtained when applied this neural network.

Contents

The interfaces of this system are designed using MATLAB graphical user interface. Login is the first interface must be appear to allow the operator enter to system, if the user name and password is

CHAPTER ONE

CHAPTER ONE

INTRODUCTION

1.1 General

Unmanned aerial vehicles (UAV) preprogrammed to do its mission and come back without human interference. Many methods are used to control UAV such as manual control using joystick, or by PC, remote control and so on. Recently android is used to control them. These methods mentioned above required the operator who controls UAV through remote or mobile has a deep knowledge about the controller and the UAV dynamic. Another simple method can be used which is voice recognition. Voice recognition is the techniques which convert the voice in to sequence of word.it is used for control the devices such as cars. Using voice recognition in controlling UAV is very easy to use and operator doesn't need information about the controller. Thus this method is proposed in this thesis to control UAV. Many software and hardware are available to apply this technique; MATLAB will be used in this research.

1.2 Problem Definition

Manual controller as the current interaction method requires the operator to be more professional of the specific controller and UAV dynamic behavior. Also controlling UAV manually is relatively less quick specially when controlling multiple UAV is required, thus more quickly and intelligent method is needed.

1.3 Objectives

- Apply one of the voice recognition algorithm
- Design and Write algorithm of controlling UAV
- Simulate controlled UAV motion

1.4 Scope

Control UAV using voice recognition by programming the At mega microprocessor installed on UAV.

1.5 Methodology

The flow chart of autopilot missions for the design of small UAVs or MAVs is shown. The code for this autopilot is to be written relying on the available UAV data. The voice is to be recorded by computer microphone and then is to be recognized using neural network in MATLAB environment .UML is to be analyze for the system .the result of recognition is to be shown graphically .finally the whole system is to be simulated to show the result of controlling

1.6 Thesis Outline

The outline of this thesis is as follows:

- Chapter one gives an introduction to the problem definition, research objectives, methodology of this thesis including general layout
- Chapter two present a background and literature review to define voice recognition, its principles. The theory for voice recognition to control electronic device, and the methods to control UAV.A review of electronic component of system are presented.

- Chapter three describes tools and techniques used in this thesis.
- Chapter four describes design steps and system simulation.
- Chapter five present simulation results and discussion.
- Chapter six gives conclusion and recommendations for further

CHAPTER TOW

CHAPTER TOW
BAKGROUND AND LITERATUR REVIEW

2.1. Introduction:

This chapter will show literature review and how to design the speech recognition system using the feature extraction there will be a review of the use Neural Networks within speech recognition, also the use of microcontroller within UAV control will be presented.

2.2. Speech recognition

Speech recognition applications are becoming more and more useful now a days. Speech is most, effective and natural medium to exchange the information among people. Speech recognition applications are becoming more and more useful now a days. Speech is most, effective and natural medium to exchange the information among people. Speech recognition applications are becoming more and more useful now a days. Speech is most, effective and natural medium to exchange the information among people.

2.3. The process of speech recognition:

Speech recognition can be defined as the ability of machine to understand the speech. There is large number of characteristic need to be considered for speech recognition:

1- Vocabulary size:
Small size of vocabulary lead to more efficient speech recognition, the amount of error increases when this size is increased, and hence the process of speech recognition becomes less efficient.

2-**Speaker dependence vs. independence:**

The difference in voices for specific word depend mainly on the speakers , to recognize speech that the system accept from different speaks is more difficult and complex and increases the amount of error.

3-**Discontinuous or continuous speech:**

Continuous speech means there is no large silence between words, and this continuous make the words interleaved and more difficult for machine to understand the words.To recognize speech there are number of steps that machine must follow these steps are shown in figure (2.1).

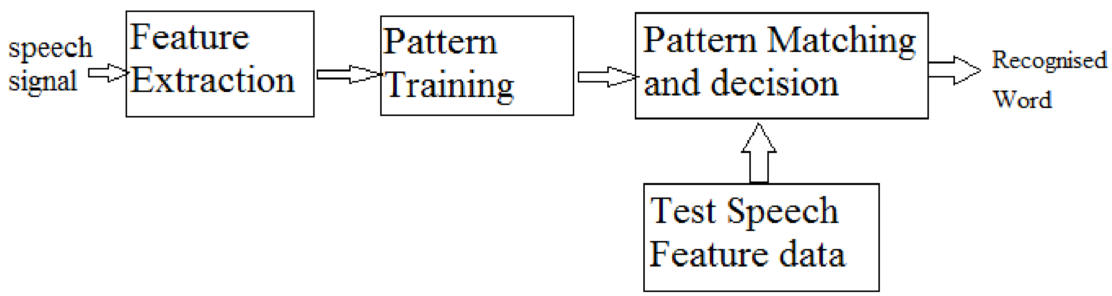


Figure 1.2: the process of speech recognition

2.4. Feature Extraction:

To recognize speech the features that distinguishes between words is needed to be known and this is the main and most important part of speech recognition which is known as Feature extraction .The goal of this part is to calculate feature vectors to provide compact representation of input signal. Feature extraction is performed in three stages. First stage is speech analysis or acoustic front end. This stage analyzes signal spectra. The second stage compiles extended feature vector composed of static and dynamic features.to transform the feature vectors into more suitable vectors to the recognizer the third stage which perform this mission is needed.[1]

2.4.1 Feature extraction using MFCC:

The Mel Frequency Cepstral Coefficients (MFCC) technique one of the many others techniques used to extract features. Tones with different frequencies that compose speech signal have a subjective pitch measured on the ‘Mel’ scale individually. To extract important characteristics of speech, signal is to be expressed in the Mel frequency scale. The Mel scale relates perceived frequency, or pitch, of a pure tone to its actual measured frequency. Small changes in pitch at low frequencies are much better at discerning by human than they are at high frequencies. Incorporating this scale makes features match more closely what humans hear. The Mel -frequency scale is a linear frequency spacing below 1000 Hz (, 40dB) which represent perceptual hearing threshold, and a logarithmic spacing above 1000Hz. Mel-frequency cepstrum (MFC) represent the short-term power spectrum of a sound, based on a linear cosine transform of a log power spectrum on a nonlinear Mel scale of frequency. The coefficients that collectively make up an MFC known as Mel-frequency cepstral coefficients (MFCCs) . The difference between the cepstrum and the Mel-frequency cepstrum is that in the MFC, the frequency bands are equally spaced on the Mel scale, which approximates the human auditory system's response more closely than the linearly-spaced frequency bands used in the normalcepstrum. This frequency warping can allow for better representation of sound [1]. To extract feature using MFCC many steps are illustrated in a block diagram of the MFCC processes shown in figure 2.2

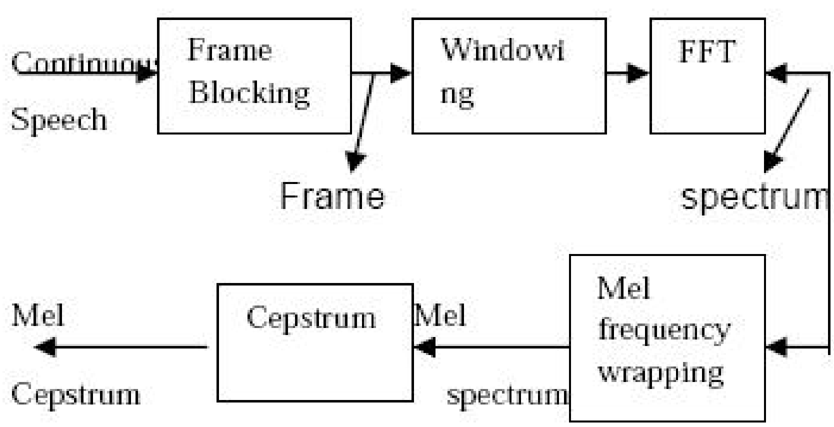


Figure 2.2: Block diagram of MFCC

Step 1: Frame Blocking:

First sound must be silence removed, to remove these silence the wave is divided into small frames of N samples with the same length which is in the range of 20 to40 msec. The first frame consists of the first N samples. The second frame begins M samples after the first frame, and overlaps it by N - M samples and so on. This process continues until all the speech is accounted for within one or more frames. Adjacent frames are being separated by: M (M<N).Typical values used are M = 100 and N= 256.figure (2.3) show the frame blocking process.

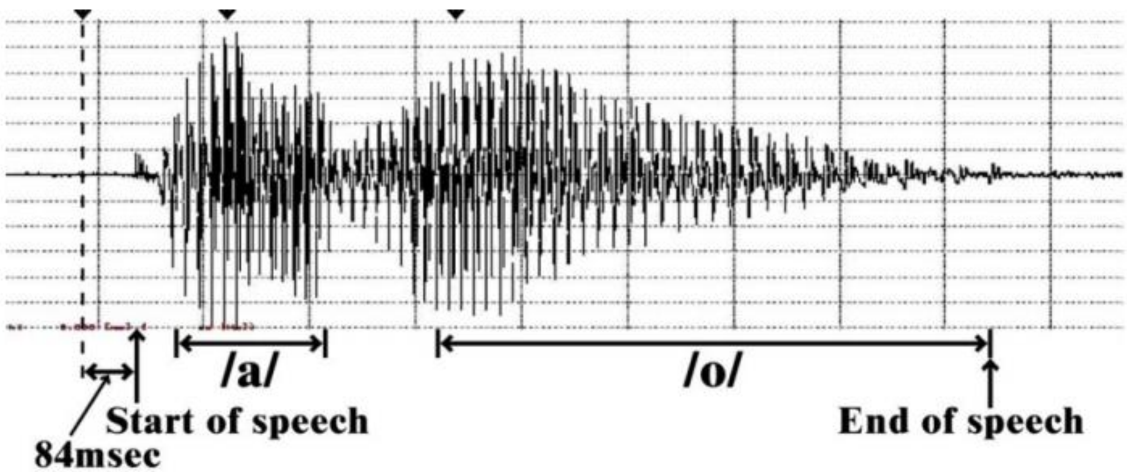


Figure 3.2: frame blocking process

Step 2: windowing step:

The goal of windowing each individual frame is to minimize the signal discontinuities at the beginning and end of each frame which means minimize the spectral distortion by tapering the signal to zero at the beginning and end of each frame

Step 3: FFT (Fast Fourier Transform):

Each frame of N samples should be represented in frequency domain by converting it from time domain to frequency domain using Discrete Fourier Transform of the windowed data to compute its spectral features. This is done using the FFT algorithm, The Fourier Transform is used to convert the convolution of the glottal pulse and the vocal tract impulse response in the time domain into frequency domain.

Step 4: Mel Filter Bank Processing:

Because voice signal does not follow the linear scale and the frequencies range in FFT spectrum is very wide the bank of filter depend on Mel scale is performed. a set of triangular filters that are used to compute a weighted sum of filter spectral components so that the output of process approximates to a Mel scale is shown in Figure.

Step 6: Cepstrum:

The log Mel spectrum should be converted back to time. The result is called the Mel Frequency Cepstrum Coefficients (MFCC).The cepstral representation of the speech spectrum provides a good representation of the local spectral properties of the signal for the given frame analysis. Because the Mel spectrum coefficients (and so their logarithm) are real numbers,so the discrete cosine transform (DCT)used to convert log Mel spectrum to time domain .

2.5. Neural network

A neural network can be defined as a model of reasoning based on the human brain. The brain consists of a densely interconnected set of nerve cells, or basic Information processing units, called neurons. The human brain incorporates Nearly 10 billion neurons and 60 trillion connections, synapses. By using multiple neurons simultaneously, the brain can perform its functions much faster than the fastest computers . A neuron consists of a cell body, soma, a number of fibres called dendrites, and a single long fibre called the axon. While dendrites branch into a network around the soma, the axon stretches out to the dendrites and somas of other neurons. [2]

2.5.1. Neuron

A neuron consists of a cell body, soma, a number of fibres called dendrites, and a single long fibre called the axon. While dendrites branch into a network around the soma, the axon stretches out to the dendrites and somas of other neurons.

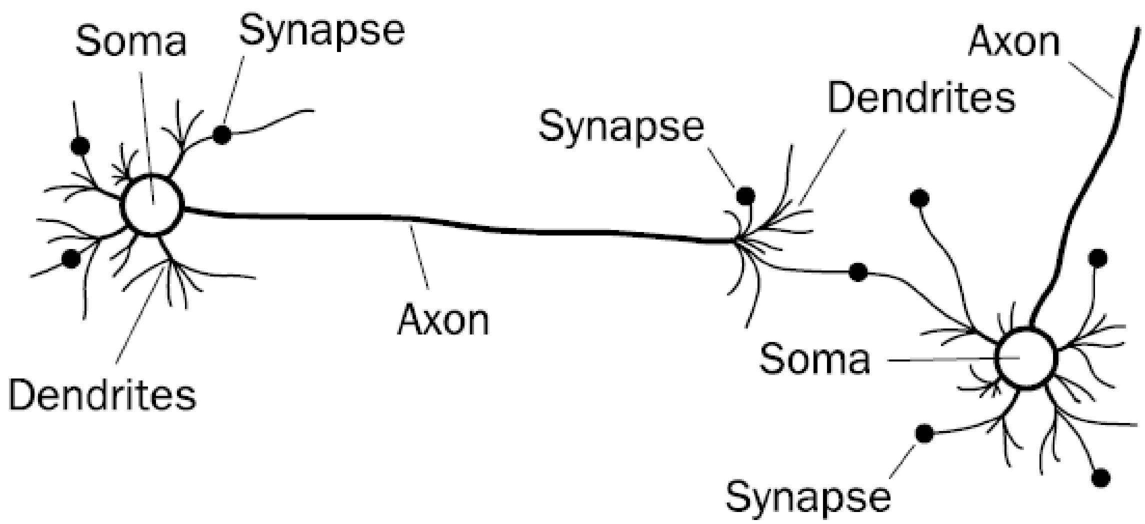


Figure 4.2: Biological neural network

A neuron receives several signals from its input links, computes a new activation Level and sends it as an output signal through the output links. The input signal Can be raw data or outputs of other neurons. The output signal can be either a Final solution to the problem or an input to other neurons. [2]

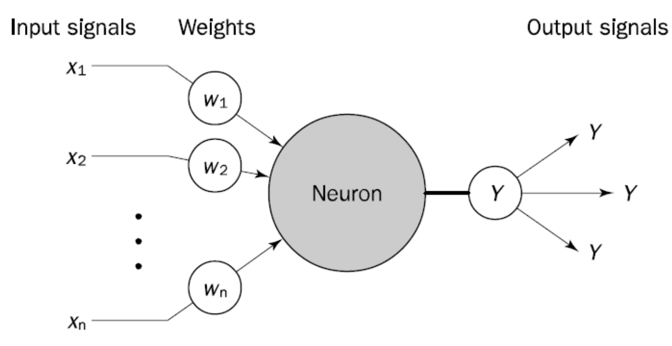


Figure 5.2: Diagram of a neuron

The neuron computes the weighted sum of the input signals and compares
The result with a threshold value, θ If the net input is less than the threshold, the Neuron output is -1. But if the net input is greater than or equal to the Threshold, the neuron becomes activated and its output attains a value +1

2.5.2. Type of activation function:

- step functions
- sign functions
- sigmoid functions
- linear functions

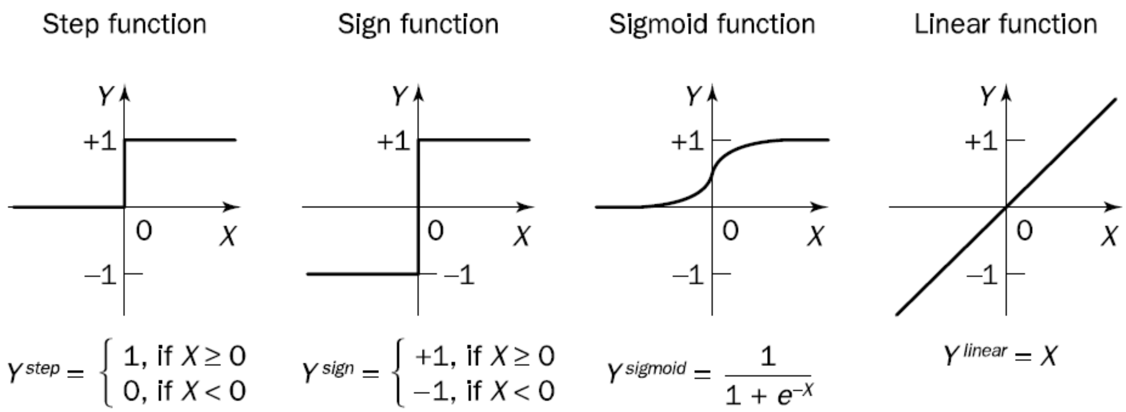


Figure 6.2: Activation functions of a neuron

2.5.3. Department of classification

- **perception**

The perceptron is the simplest form of a neural network. It consists of a single neuron With adjustable synaptic weights and a hard limiter

- **Multilayer neural networks**

A multilayer perceptron is a feed forward neural network with one or more Hidden layers. Typically, the network consists of an input layer of source Neurons, at least one middle or hidden layer of computational neurons, and An output layer of computational neurons. The input signals are propagated in a Forward direction on a layer-by-layer basis.

In a perceptron, there is only one weight for each input and only one output.

But in the multilayer network, there are many weights, each of which contributes

To more than one output. [2]

2.5.4. Learning Artificial Neural Network

The neurons are connected by links, and each link has a numerical weight Associated with it. Weights are the basic means of long-term memory in ANNs. They express the strength, or in other words importance, of each neuron input. A neural network 'learns' through repeated adjustments of these weights. [2]

2.6. Servo Motor

A DC motor is equipped with an electronic circuit to control precisely the direction of the motor shaft rotation and angle, And a very large torque, is used frequently in motion systems in the machinery of motorized robots, moving radar devices and in moving aircraft wings.

Figure 7.2 servo motor

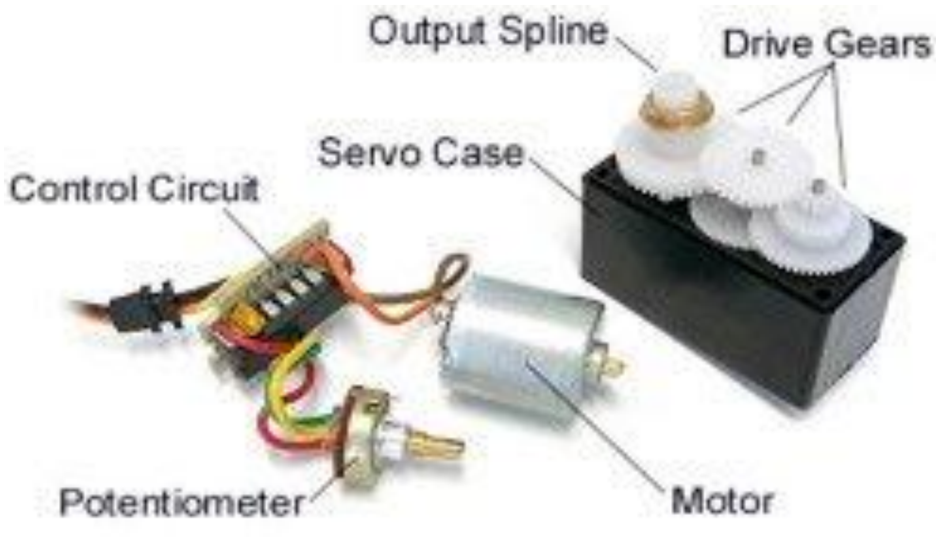
2.6.1.It consists of

1. DC Motor Its function is to generate the motion that is needed to move the gears.

2. Resistant variable function Give the value of the voltage corresponding to the position of the torque motor based on the value of its resistance as it moves with the column of the engine rotation, the column is the column of the column which changed its resistance value.

3. Gears function shift torque motor speed.

4. Circuit electrically function receives incoming signal of variable resistance.



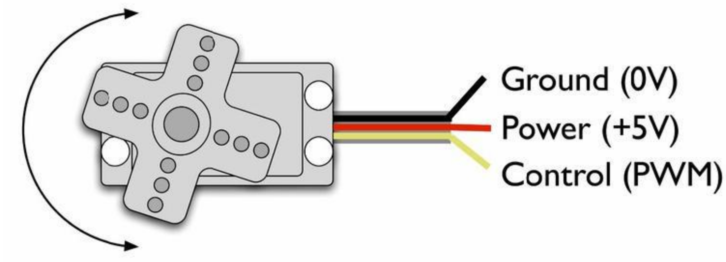
-

Figure 8.2: component of the servo

2.6.2. Control of servomotor

Most servomotor use a stander three wire interface:

- Power
- Ground
- Control line



2.6.3.How to work

-To turn the servo motor in a certain direction that needs a 50Hz pulse range and pulse width HIGH + LOW = 20ms

-For a clockwise rotation, we give it a pulse of 1 (1.0: 0.7) milliseconds and a pulse of 0 (19) milliseconds. The periodic time is 19 + 1 or 20 Ms.

-To turn counterclockwise, we give it a pulse of 1 (2.0: 1.7) Ms. and a pulse of 0 (18) milliseconds. The periodic time is 18 + 2, i.e., 20 Ms.

-To return to the total position of 0 °, we give the pulse 1 for 1.5 Ms. and 0 pulses for 18.5 Ms. the periodic time is 18.5 + 1.5 MS 20 m .[3]

2.7. Microprocessor and microcontroller

2.6.1 What the Microprocessor

Is a computer electronic component made from miniaturized transistors and other circuit element on a single semiconductor integrated circuit (IC). The central processing unit (CPU) is the most well-known microprocessor but many other components in a computer have them such As the Graphics Processing Unit (GPU) On a video card. It is a computer processor on a microchip and is a multipurpose, programmable device that uses digital data as input and provides results as an output once it processes the input according to instructions stored in its memory, they are designed to perform arithmetic and logic operations that make use of data on the chip. General purpose microprocessors in PCs are used for multimedia display, computation, text editing and communication. Several microprocessors are part of embedded systems. These embedded microprocessors provide digital control to several objects including appliances, automobiles, mobile phones and industrial process control. [4]

Figure 9.2: Microprocessor

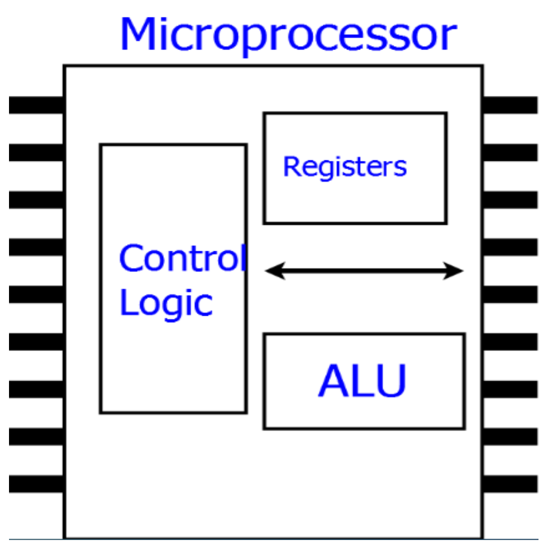
2.6.2 What the microcontroller

A microcontroller is a computer present in a single integrated circuit which is dedicated to perform one task and execute one specific application. It contains memory, programmable input/output peripherals as well a processor. Microcontrollers are mostly designed for embedded applications and are heavily used in automatically controlled electronic devices such as cellphones, cameras, microwave ovens, washing machines, etc. [5]

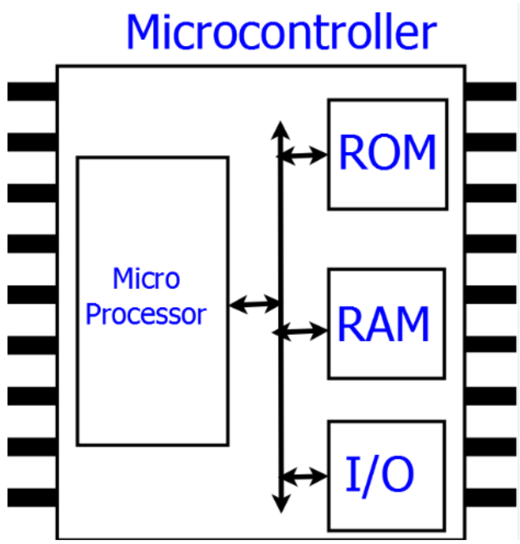
Figure 10.2: microcontroller

2.6.3 What is the difference between a microprocessor and microcontroller?

A microprocessor is an Integrated Circuit (IC) which has only the Central Processing Unit (CPU) inside it. They lack Random Access Memory (RAM), Read Only Memory (ROM), and other peripherals on the chip. To deploy a microprocessor in a system, the above-mentioned peripherals are required to be connected. The Intel Pentium series, i3, i5, Cortex A8 are popular microprocessors which find their applications in desktop PC's, Laptops, notepads etc.



Unlike microprocessors, microcontrollers possess a CPU along with RAM, ROM, and other peripherals on the same chip. Therefore, it is also termed as a mini computer. Though the size of RAM and ROM and the processing power is far less as compared to an actual computer, it meets all the characteristics of a complete computer. Popular microcontrollers in the market are 8051, STM32, PIC32, Arduino, ATMEL etc. [6]



2.8. Literature review

1.Seemaramdhaf and shilpajoshi designed robot controlled by android device through the speech recognition, the speech synthesis would be done by smartphone android app and send the recognized commands to the robot via Bluetooth which would be placed on the robot and then these command would be analyzed by the Microcontroller on the robot which would control the robot using motor driver [7].

Summary: the thesis designed robot controlled by android device through the speech recognition and send the recognized commands to the robot via Bluetooth.

are used in this thesis send the recognized command to UAV through the GSM.

2. Safaa Omer Mohammed is classification model for four different Altajweed rules like the Allah name (mofakham , morakaq) and moon and sun L (لام) as we depended on two different kinds of voice features LPC(liner predictive coding) and MFCC(Mel-Frequency cepstrum), and depended on two classifying mechanisms(Neural Networks and Hidden Markov Model(HMM)).

Result: In the training of Hidden Markov Models accuracy amount 90% with Allah (moufakhum), 92% with allah (mourqeq), 83.3% with sunny لام and 80% with moony لام. In Neural Network able to distinguish between four rules of tajweed and training samples processing with high accuracy reached

95% with Allah (moufakum), 94% with Allah (mourqeq) 93% with moony لام and 92.3% with sunny لام. [8]

Summary: they use LPC , MFCC , Neural Networks and Hidden Markov Model are used in this thesis the speech features MFCC and classifying mechanisms Neural Networks

3-Maye Abdalslam and Hiba Abdall Rahman controlled a wireless car remotely by computer. The system was designed by adding some simple electronic components to the remote control and connecting it to the computer via the parallel port. The user can drive and guide the vehicle through a simple interface, This control includes a control of the speed of the car and control the amount of rotation of the wheels at different angles and each control of its own resistance is installed in the electronic circuit through which the control of the car, the use of Matlab technology to process inputs and control the car accordingly.[9]

Summary: controlled in wireless car through remotely by computer, are used in this thesis control UAV through the computer using simulation and speech recognition

4. Abu-Bakr and Osama, Mortda controlled of wireless air plane which controlling by remote control device into controlling using computer.

Install a wireless camera to track movement of the wireless aircraft, in addition to the possibility of dropping things form the plane. Software java, C language and assembly language and then has used (JNI) technique to linking these language, the program also was used MATLAB to deal with the photos and videos. [10]

Summary: controlled in wireless air plane through remotely by computer, are used in this thesis control UAV through the computer using simulation and speech recognition

5- Ashraf Mohammed select five verses from the Holy Quran to implement the voice recognition using three techniques of speech recognition are used. They are Hidden Markova Models. Dynamic time warping and Artificial Neural Networks .they are used Mel frequency Cepstrum Coefficients MFCC .Power spectral density PSD and Reflection Coefficients RC to extract the feature from the speech.

Result: HMMs scored high rates of recognition to coefficients used (MFCC, PSD and RC) low recognition rates with high confusion scored by ANNs and DTW at verse level. For all coefficients used, high scores of recognition rates with low confusion rate concentrated in HMMs with MFCC. MFCC Scored higher recognition rates than PSD and RC.[11]

Summary: the train the neural network to recognize different voice. are used in this thesis train the neural network to recognize one voice Alexander Murphy Implementing Speech Recognition with Artificial Neural Networks use 5 sample trains and 5 sample test. Use MFCC to extract the feature from the speech

Result: the results were still very positive. Despite limiting the speech recognition side of the project, I gained an understanding of how neural networks can tackle a problem like pattern recognition, as well as the benefits of certain structures and training algorithms. In the end, I accomplished what I set out to and successfully implemented speech recognition with a neural network.[12]

Summary: use 5 sample trains and 5 sample test and not implement it . 15 sample trains and 5 sample test are used in this thesis and have been implemented to control the UAV .

CHAPTER THREE

CHAPTER THREE

TECHNIQUES AND UML

3.1. Introduction:

In this chapter we will describe the software tools used in this project and analysis the project with UML tools.

3.2. Matlab Package:

MATLAB ("Matrix Laboratory") is a tool for numerical computation and visualization. The basic data element is a matrix, so if you need a program that manipulates array-based data it is generally fast to write and run in MATLAB (unless you have very large arrays or lots of computations, in which case you're better off using C or Fortran.

MATLAB has many advantages compared to conventional computer languages (e.g., C, FORTRAN) for solving technical problems. MATLAB is an interactive system whose basic data element is an array that does not require dimensioning. The software package has been commercially available since 1984 and is now considered as a standard tool at most universities and industries worldwide.

There are Specific applications are collected in packages referred to as toolbox. There are toolboxes for signal processing, symbolic computation, control theory, simulation, optimization, and several other fields of applied science and engineering.

In this research we used Mat lab Package in signal processing, in implement neural networks to recognize speech, interfaces and simulation.

3.2.1. Features of MATLAB:

- It is a high-level language for numerical computation, visualization and application development.

- It also provides an interactive environment for iterative exploration, design and problem solving.
- It provides vast library of mathematical functions for linear algebra, statistics, Fourier analysis, filtering, optimization, numerical integration and solving ordinary differential equations.
- It provides built-in graphics for visualizing data and tools for creating custom plots.
- MATLAB's programming interface gives development tools for improving code quality maintainability and maximizing performance.
- It provides tools for building applications with custom graphical interfaces.
- It provides functions for integrating MATLAB based algorithms with external applications and languages such as C, Java, .NET and Microsoft Excel.

3.2.2. Neural Network:

Neural Network Toolbox™ provides algorithms, retrained models, and apps to create, train, visualize, and simulate both shallow and deep neural networks. You can perform classification, regression, clustering, dimensionality reduction, time-series forecasting, and dynamic system modeling and control.

3.2.3. Interfaces:

GUIs (also known as graphical user interfaces or UIs) provide point-and-click control of software applications, eliminating the need to learn a language or type commands in order to run the application.

3.2.4. 3D animation

Simulink 3D Animation™ provides apps for linking Simulink® models and MATLAB® algorithms to 3D graphics objects. It lets you visualize and verify dynamic system behavior in a virtual reality environment. Objects are represented in the Virtual Reality Modeling Language (VRML), a standard 3D modeling language. You can animate a 3D world by changing position, rotation, scale, and other object properties during desktop or real-time simulation. You can also inject virtual sensor signals and access 3D animation data in Simulink or MATLAB for post processing.

Simulink 3D Animation includes viewers for rendering and interacting with virtual scenes. With the 3D World Editor, you can author detailed scenes assembled from 3D models exported from CAD-based or web-based sources. You can incorporate multiple 3D scene views inside MATLAB figures and interact with these views via a force-feedback joystick, space mouse, or other Hardwar.

In this project we used the 3d animation to Aircraft Simulation.

3.3. Enterprise:

Graphical tools design help you to trace-high level specification to analysis ,design , implementation ,test and maintenance models using UML , and other open standards. We use the UML standards to analysis the project.

3.4. Proteus:

Is a proprietary software tool suite used primarily for electronic design automation the software is used mainly by electronic design engineers and electronic technicians to create electronic schematics and electronic prints for manufacturing printed circuit boards. In this project we used Proteus to Microcontroller simulation. The microcontroller simulation in Proteus works by applying either a hex file or a debug file to the microcontroller part on the schematic.

3.5. Micro AVR

Micro PRO for AVR is a full-featured ANSI C compiler for AVR devices from Atmel. It features intuitive IDE, powerful compiler with advanced SSA optimizations, lots of hardware and software libraries, and additional tools. Compiler comes with comprehensive Help file and lots of ready-to-use examples designed to get you started in no time. Compiler license includes free upgrades and a product lifetime tech support, so you can rely on our help while developing.

3.6. ATmega16

In this project use the atmega16 .ATmega16 is an 8-bit high performance microcontroller of Atmel's Mega AVR family with low power consumption. Atmega16 is based on enhanced RISC (Reduced Instruction Set Computing, Know more about RISC and CISC Architecture) architecture with 131 powerful instructions. Most of the instructions execute in one machine cycle. Atmega16 can work on a maximum frequency of 16MHz.ATmega16 has 16 KB programmable flash memory, static RAM of 1 KB and EEPROM of 512 Bytes. The endurance cycle of flash memory and EEPROM is 10,000 and 100,000, respectively. ATmega16 is a 40 pin microcontroller. There are 32 I/O (input/output) lines which are divided into four 8-bit ports designated as PORTA, PORTB, PORTC and PORTD.

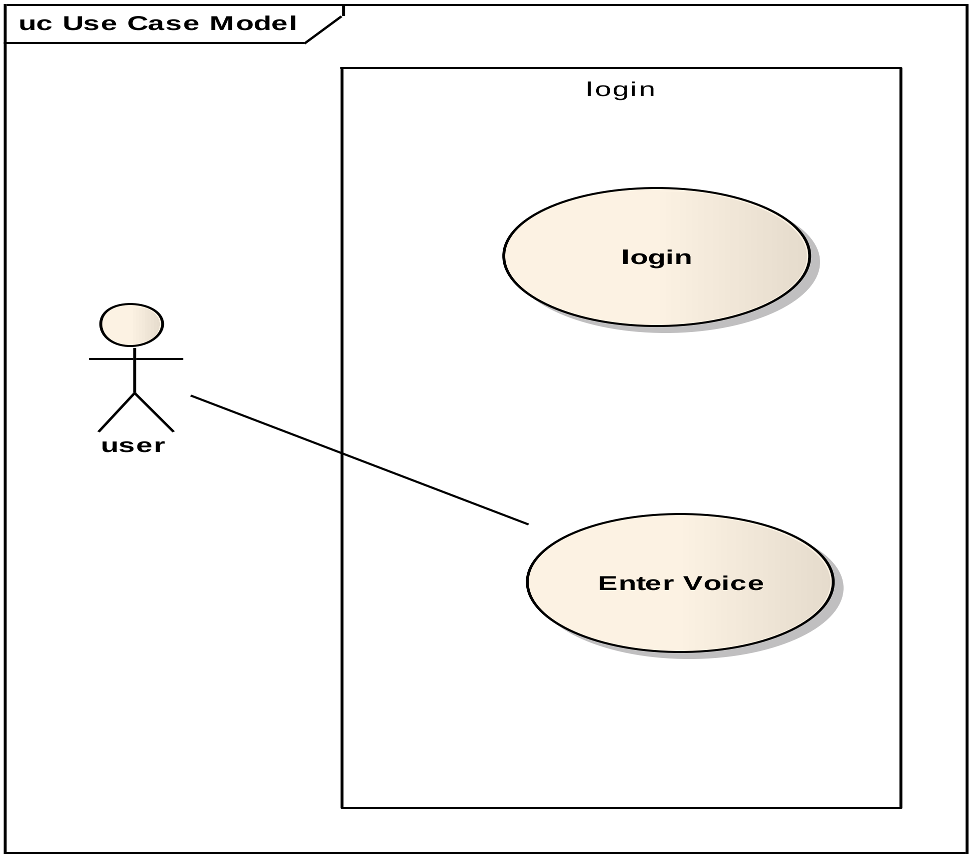


Figure3. 1 use case Model of the system

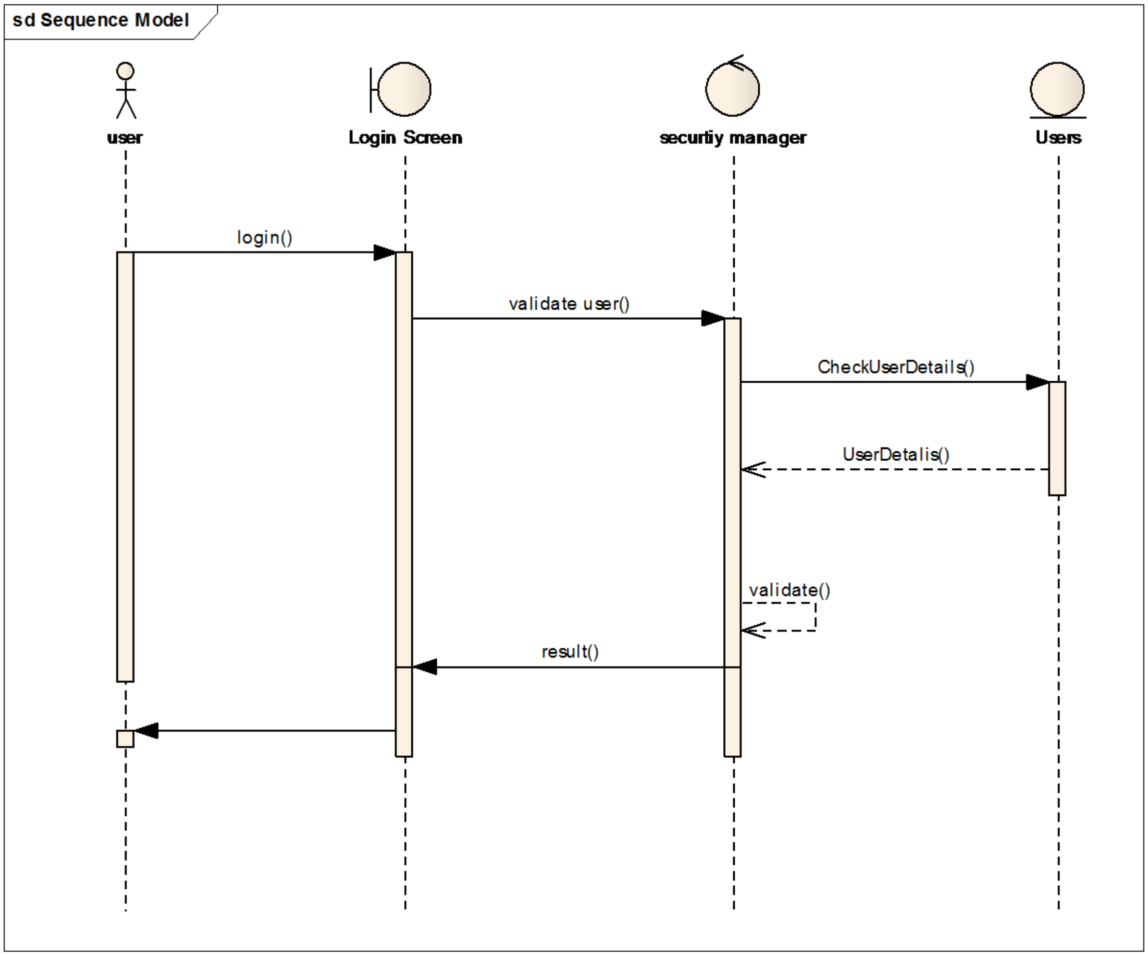


Figure3. 2: Sequences Model of the login screen

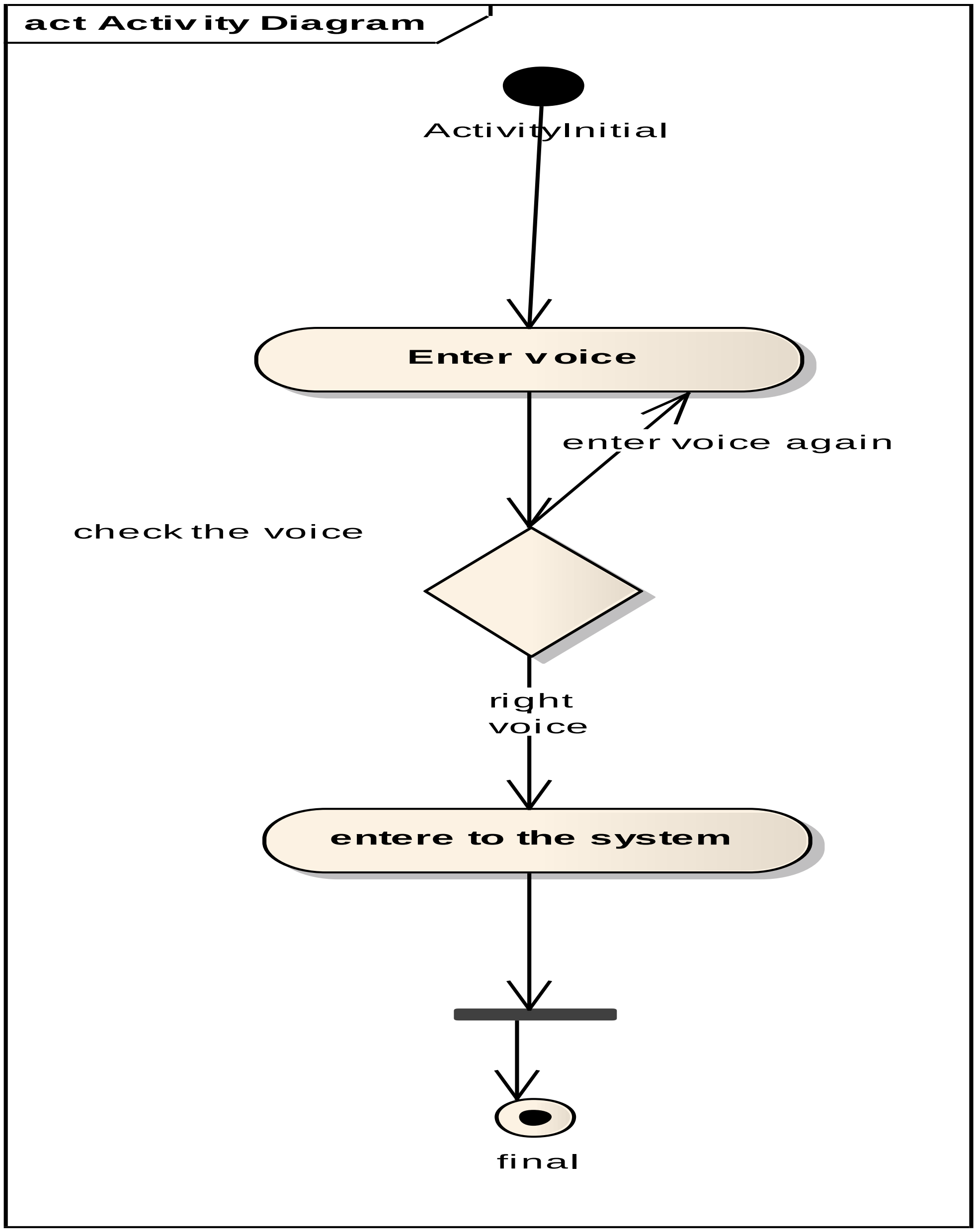


Figure3. 3: Activity Model of the Enter voice

CHAPTER FOUR

CHAPTER FOUR

DESIGN AND SIMULATION

4.1 Introduction

The theory of recognized the four specified word will be described as well as process the audio samples and in another side explain how the UAV can be controlled by these recognized word ,the simulation and coding controlled UAV motion are also be introduced in this chapter.

4.2 Design Requirements

1. Error to be less than 10% for recognizing word by neural network
2. Arrive to the right motion of UAV according to specific word
3. The system will be speaker dependence and will recognize the voice of only one user

4.3 Audio Processing

To recognize patterns using neural network, features distinguishes between these patterns must be extracted and inserted to neural network as input , commonly used features for speech and audio samples are mel-frequency cepstral coefficients, in order to extract these features there are steps have been followed:

1. First the audio sample has been recorded by MATLAB sampling frequency called Fs and a duration called Duration are defined as shown in figure 4.1.

```
function [up]=rcrd
Fs=8000
nbit=8
up = audiorecorder(Fs,nbit,1)

% Define callbacks to show when
% recording starts and completes.
up.StartFcn = 'msgbox(''Start speaking.'')'
record(up, 5)
up.StopFcn = 'msgbox(''End of recording.'')'
```

Figure 4. 1:Audio recording using MATLAB

First 30 samples are recorded for each word using this MATLAB tools but later only 20 samples are taken because the limitation of hardware, the recording of these samples produce audio objects, and these objects are then converted to numbering signal represent amplitude using *get audio data* function in MATLAB. Figure 4.2 shows this conversion.

```
m=input('enter audio file')
  fs=8000
    voice = getaudiodata(m);
```

Figure 4. 2: Audio object conversion using MATLAB

2. Then each sound has been divided into small frame each frame has 0.3s of length.

3. Silence is removed from all recorded sample to reduce the size of sound and hence reduce calculation. This process has been applied to each frame. The code implement this method is obtained in Appendix A, figure 4.3 and 4.4 show the sound for '*down*' word before and after remove silence respectively.

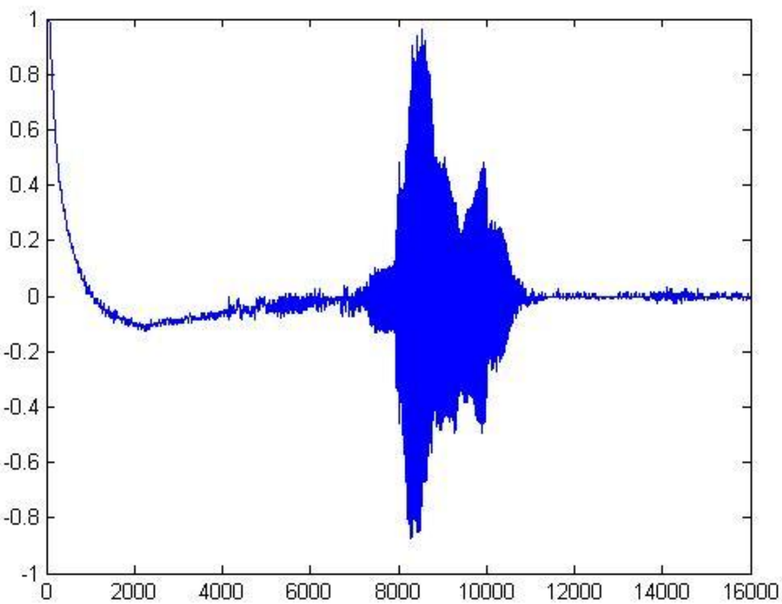


Figure 4. 3: recorded sound for ‘down’ word before removing silence

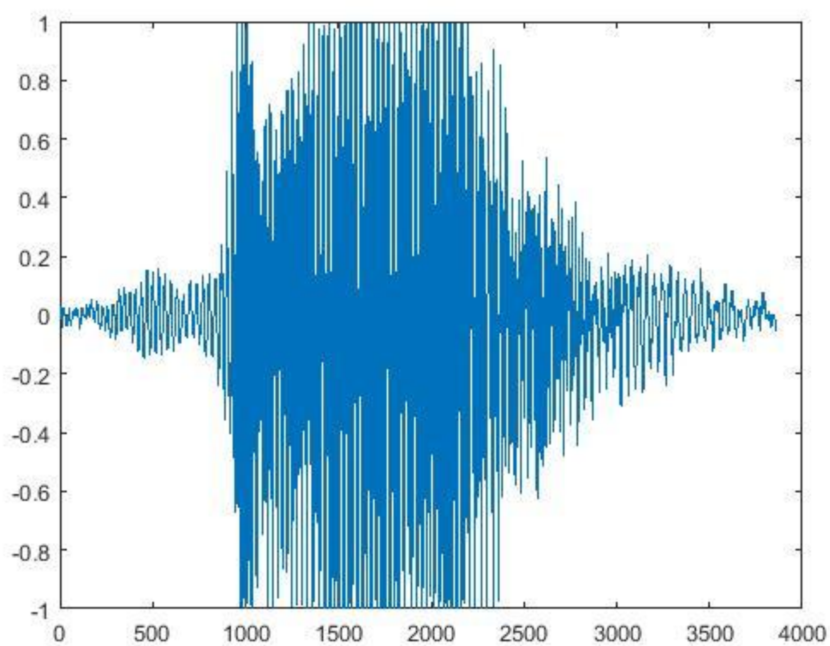


Figure 4. 4: recorded sound for ‘down’ word after removing silence

4. All signals must be at the same length to avoid 0’s and then infinity after applying MFCC function, thus the longest signal is corrupted to verify this goal.

5. Windowing each individual frame to minimize the signal discontinuities has been applied.

6. Discrete Fourier Transform has been taken of the windowed data using FFT to covert signal from time domain to frequency domain.

7. The Power spectrum then is computed.

8. Mel Filter Bank Processing and its logarithm is then performed to compute log Mel spectrum

9. The discrete cosine transform (DCT) is done forming the Mel coefficients back to time domain and the result is Mel frequency cepstrum coefficients (MFCC). The cod of this function consist of all above steps which is shown in appendix A. Figure 4.5, 4.6, 4.7, 4.8 shows the MFCC for the four words.

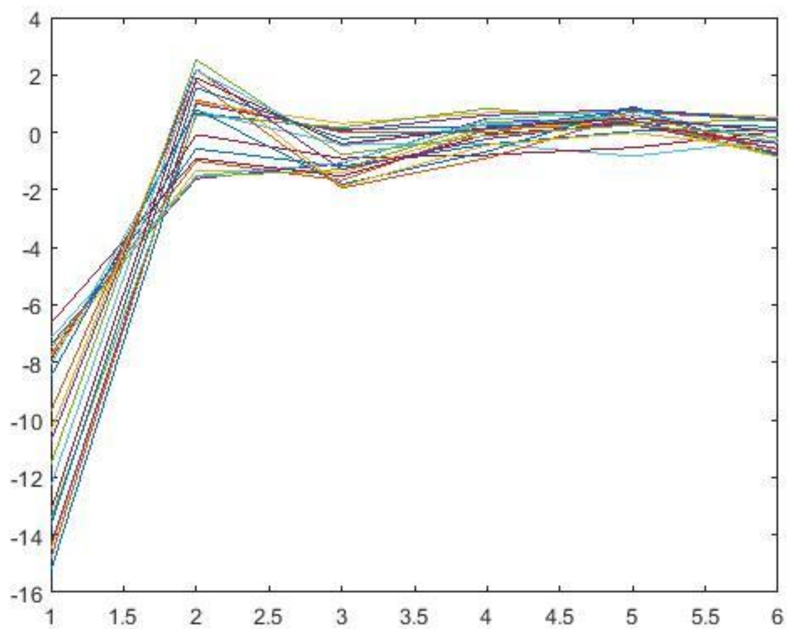


Figure 4. 5: MFCC for the word ‘down’

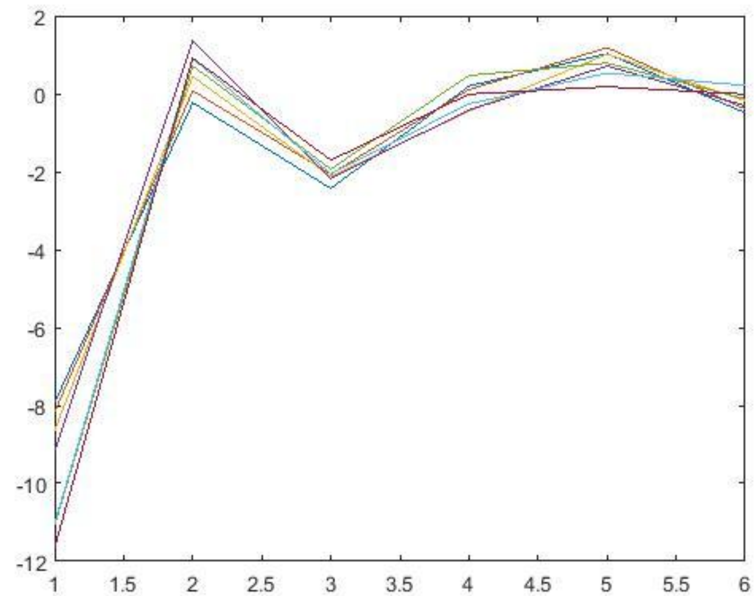


Figure 4. 6: MFCC for the word ‘up’

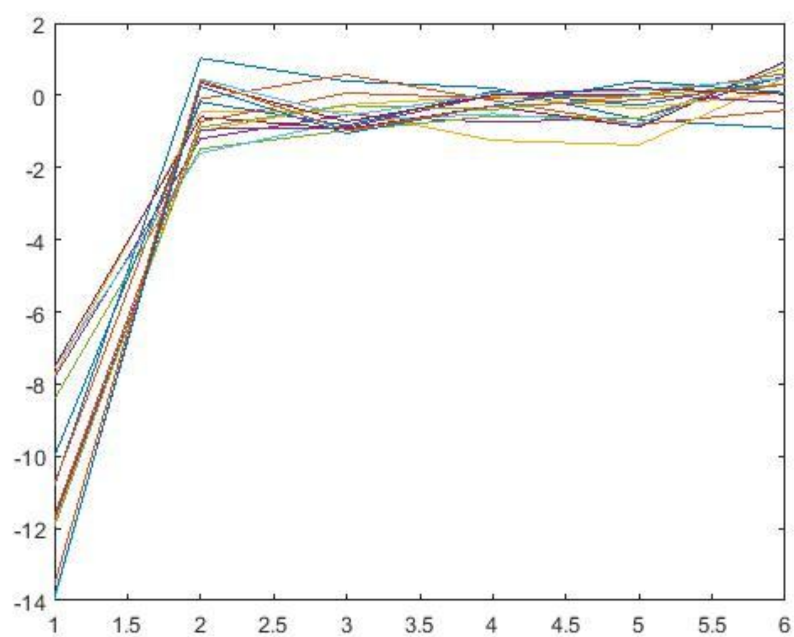


Figure 4. 7: MFCC for the word ‘left’

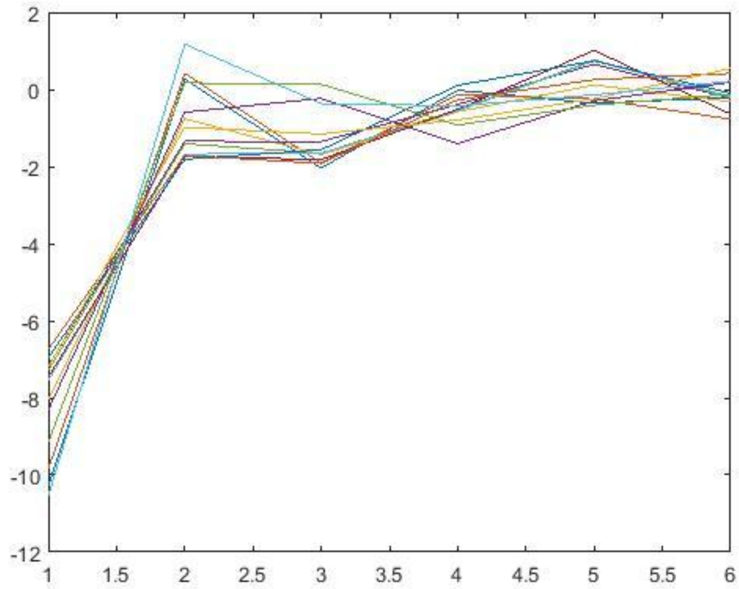


Figure 4. 8: MFCC for the word 'right'

This feature is to be inserted to neural network by forming a matrix of size 24 x 80, for training matrix which is 24 features for 80 words it resulting from taking MFCC for each word and combining them in to one matrix. this matrix is then divided into train matrix with size 24 x 60 used to train the network and test matrix with size 24 x 20 used to test to network

4.4 Structure of Neural Network

The neural network implemented in this thesis being limited to 4 specific words with four classes those are actually needed. The outputs of this network are the number of classes that network will recognize and in this case will be 4 classes, a class for each word to be recognized. The inputs are the features of each word individually that are taken from the Matlab function MFCC (), which produce 24 features for each word. The optimal number of hidden neurons can be determined by 3 common rule-of-thumb approaches:

- The number of hidden neurons should be less than twice the size of input layer
- The number of hidden neurons should be 2/3 the size of input layer plus the size of the output layer

- The number of hidden neurons should be between the size of input layer and the size of output layer.

This thesis applied the 20 neurons for training network. Figure 4.6 shows the structure of neural network implemented in this thesis.

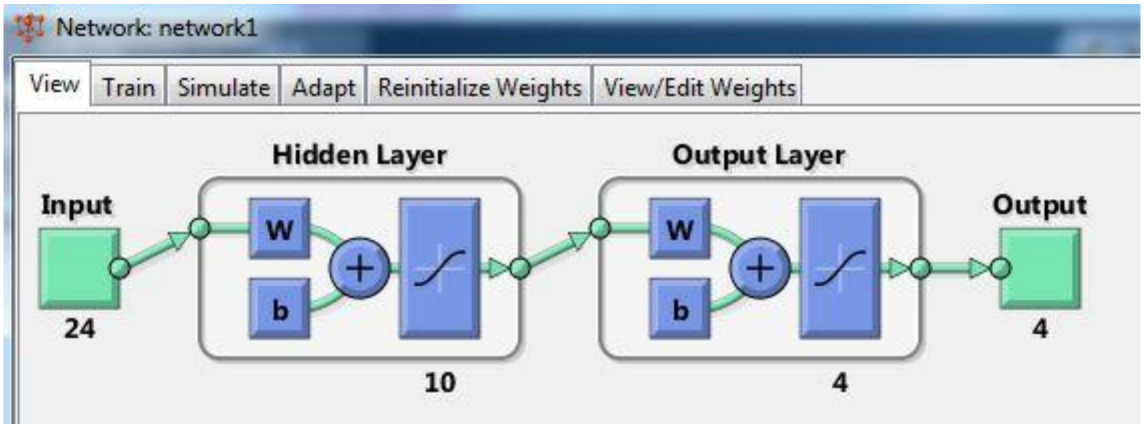


Figure 4. 9: neural network structure

Figure 4.6 shows that input consist of 24 nodes, 2 layers of 10 neurons, and an output layer of 4 representing 4 classifications of words. Back propagation is beneficial to use for pattern recognition, this algorithm is applied in this research and MATLAB neural network toolbox was used to implement this neural network. Figure 4.7 show the algorithm and function used for creating neural network and figure 4.8 shows the training of this neural network.

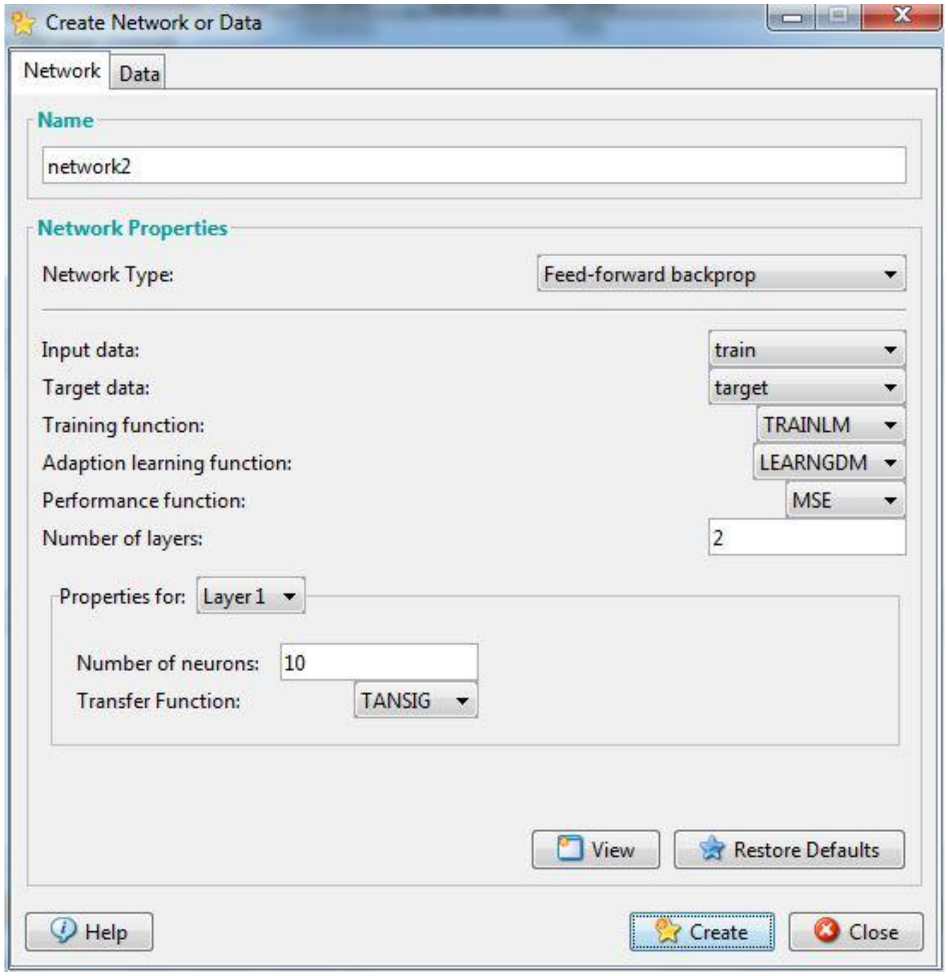


Figure 4. 10: Creating neural network

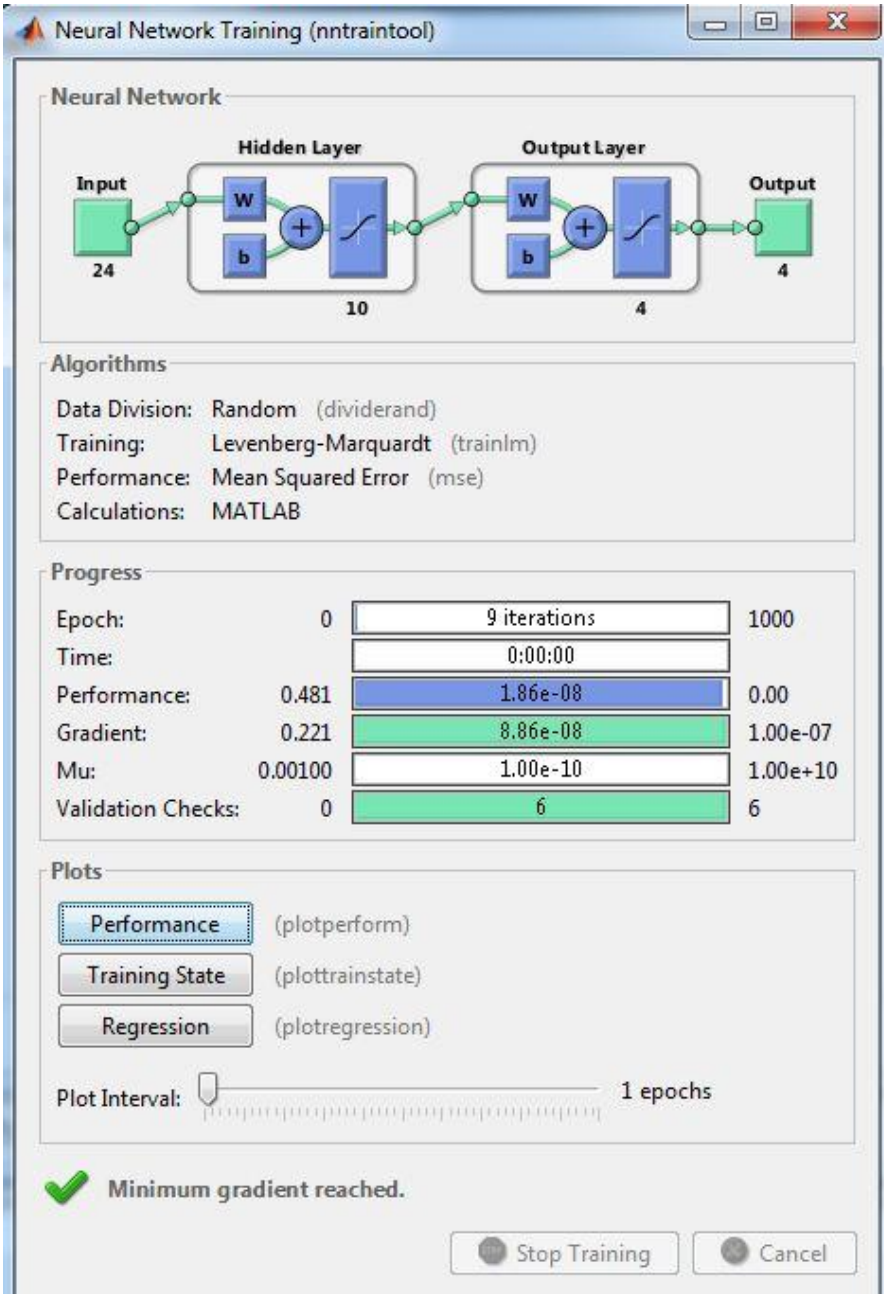


Figure 4. 11: Training neural network

To test the neural network recognizing capability the MATLAB code was written and is shown in figure 4.9.

```
function [y] = testNetwork( testSoundFile , myNetwork)
%UNTITLED2 Summary of this function goes here
% Detailed explanation goes here
z=zeros(1140,1)
for i=1:1140
z(i) = testSoundFile(i);
end
myNetwork1 = myNetwork;
%testSound = audioread(fileName);
TestSound = mfcc1(z);

ResultMatrix = sim(myNetwork1, TestSound);
Class1 = (ResultMatrix)
Class1 = round(ResultMatrix)
if Class1(1) == 1
    %for down y=1
     display('down')
    y=1

elseif Class1(2) == 1
    %for left y=2
     display('left')
    y=2
elseif Class1(3) == 1
    %for right y=3
     display('right')
     y=3
elseif Class1(4) == 1
 %for up y=4
  display('up')
y=4

end
```

Figure 4. 12: Test neural network function

The output of this function named ‘y’ takes the value of 1, 2, 3, or 4 represent the words down, left, right, and up respectively. This output was then sent to microcontroller via serial port that connects MATLAB and PROTUES.

4.5 Controlling control surfaces servo motors

ATmega 16 microcontroller is used in this thesis to control 5 servo motors , 2 servo motors are connected to elevator of UAV , the other 2 servo motors are connected to aileron ,and the final one is connected to rudder, these servo motors guided the UAV and control its motion. The theory of controlling UAV explained previously in chapter two. The Microcontroller controls these motors according to numbers which is received by serial port transmitted from MATLAB. If the number that microcontroller is received equal 1 then it moves two elevator motors 45 ° downward by sending pulse with width 1µs, and UAV moves down accordingly. These motors move 45° upward by microcontroller sending pulse with width 2 µs if it receives 2, and UAV moves up. For moving UAV left servo of the left aileron must be moving 45° downward by sending to it 1µs pulse width and the right aileron servo must be moving 45° upward by sending to it 2µs pulse width, rudder servo motor must be moved to the right 45° by receives 1 µs pulse width, this done if the microcontroller receives 3from serial port, and if it receives 4 this operation is then reversed and the UAV moving to the right. The motion of servo motors and the process of microcontroller receiving are simulated in PROTUES and the angel of these servo motors is again sent to MATLAB for simulating the motion of UAV. Figure 4.10 shows how the microcontroller is connected to serial port and servo motors

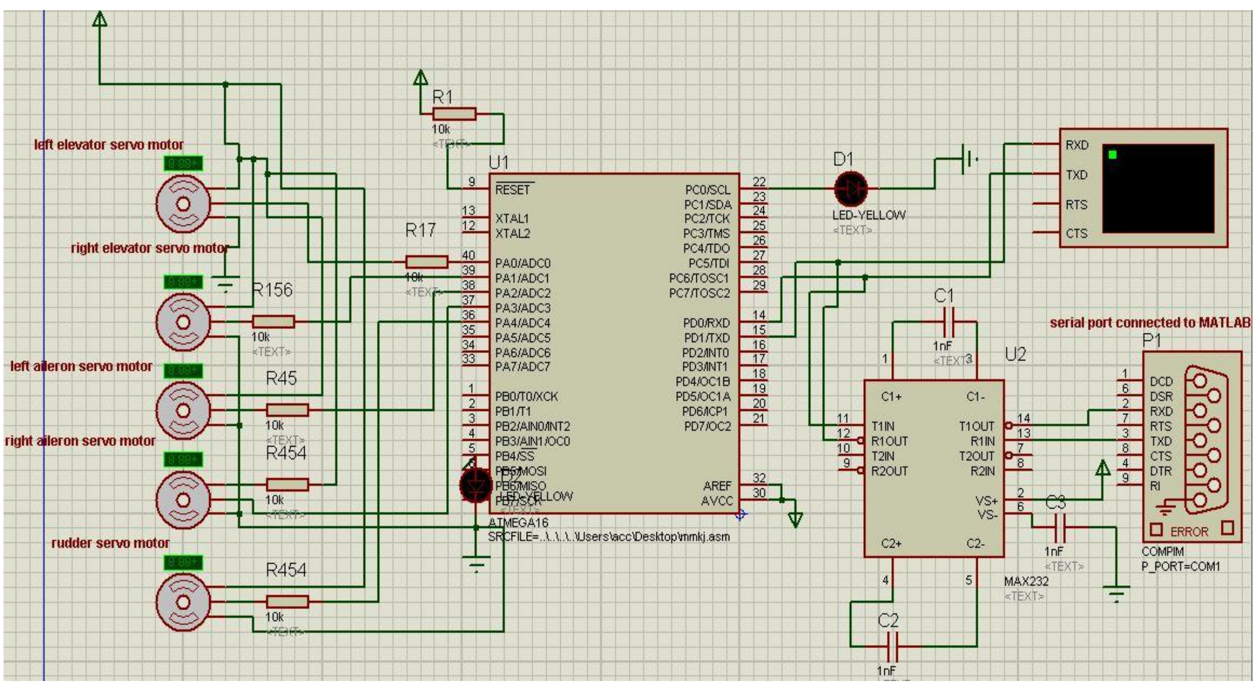


Figure 4. 13: Simulation servo motors motion and sending and receiving via serial port

The simulation of UAV is applied in MATLAB using Simulink and 3D animation tools. The function which is move the UAV to the appropriate position and attitude is written in Simulink as shown in figure 4.11 and the output of Simulink is sent to VR sink and displayed in 3D world as shown in figure 4.12.And figure 4.13 shows the 3D word editor.

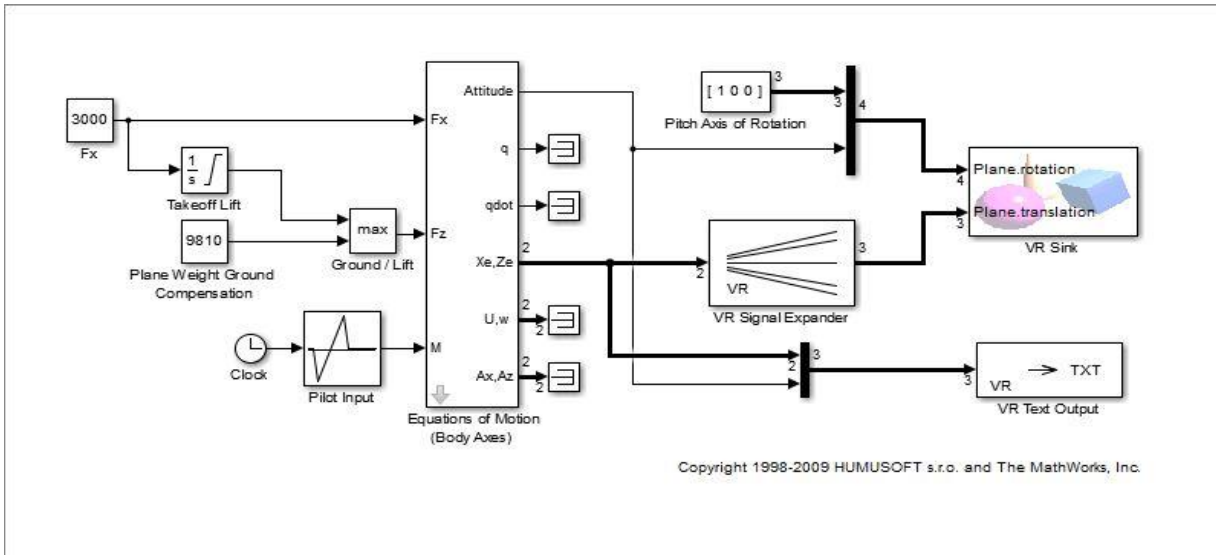


Figure 4. 14: Block diagram of simulating UAV motion

Figure 4. 15: 3D animation world for UAV motion

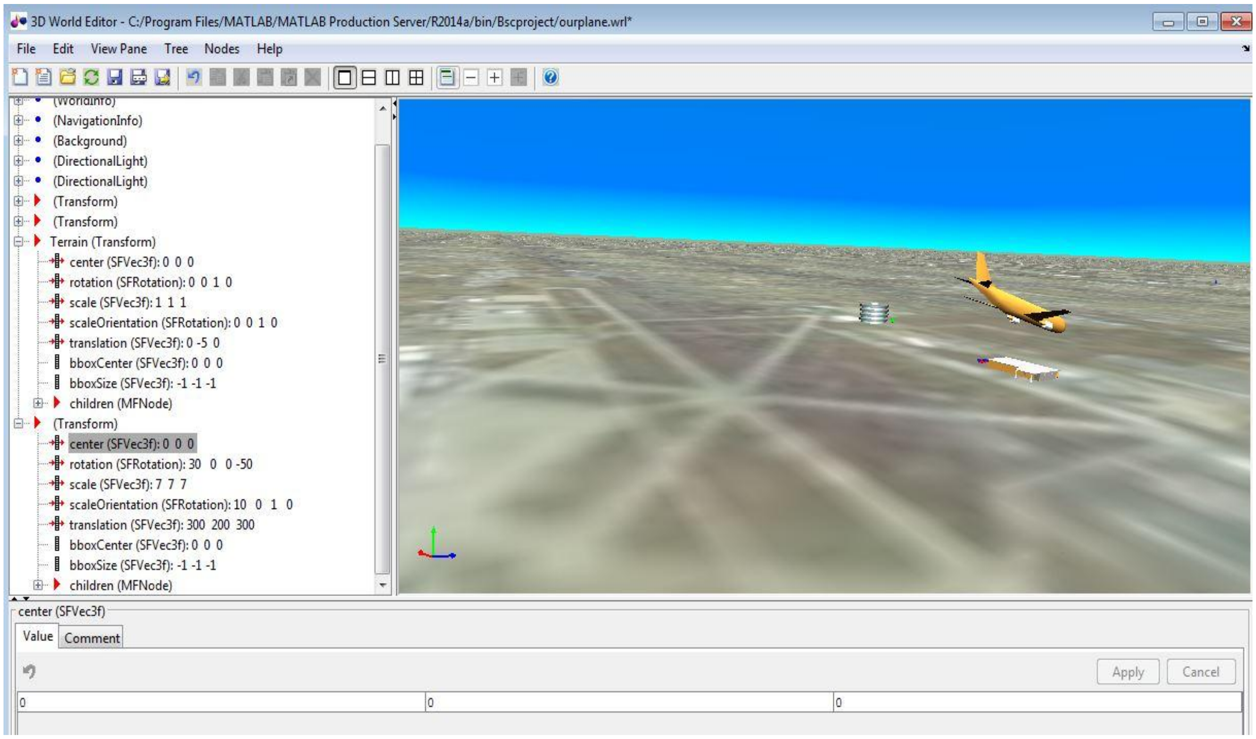


Figure 4. 16: 3D animation world editor

4.6 Design interfaces of system

The interfaces of this system are designed using MATLAB graphical user interface. Login is the first interface must be appear to allow the operator enter to system, if the user name and password is correct the second interface which allow user to enter voice command must be appear, if the system recognized the command the other function is executed and the UAV move accordingly. If the system can't recognize the command then message 'unknown words please insert your voice' must be appear. In the case of entering incorrect user name or password the message 'incorrect user name or password' will appear.

CHAPTER FIVE

CHAPTER FIVE

RESULT AND DISCUSSION

5.1 Result of neural network

The satisfactory result is obtained after training neural network with 24 features for each word. And the network well recognized these 4 words each word has 15 sample for train and 5 sample for test. Figures 5.1,5.2,5.3,5.4 and 5.5 show the target matrix of network .The training outputs of the network is shown in figures 5.6,5.7 ,5.8,5.9 and 5.10.testing matrix consist of 5 samples of each word the first five sample for ‘down’ ,the second 5 samples for ‘left ‘ ,the third samples for ‘right’ and the last five samples for ‘up’. Figure 5.11and 5.12 show the division of these words depending on target matrix. First five columns in the first row represent ‘down ‘. In the second row the five 1s represent ‘left’, and in the third row 1s represent ‘right .1s in the last row represent ‘up’. The result of testing network is shown in figures 5.13 and 5.15 represented in five columns and four rows. Figure 5.14 and 5.15 show the confusion matrix and performance of the network. Figure 5.16 shows the result after entering voice with ‘down’ word.

1	2	3	4	5	6	7	8	9	10	11	12	13
1	1	1	1	1	1	1	1	1	1	1	1	1
0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0

Figure 5. 1: First 13 value of target matrix

14	15	16	17	18	19	20	21	22	23	24	25	26
1	1	0	0	0	0	0	0	0	0	0	0	0
0	0	1	1	1	1	1	1	1	1	1	1	1
0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0

Figure 5. 2: The next 13 value of target matrix

27	28	29	30	31	32	33	34	35	36	37	38	39
0	0	0	0	0	0	0	0	0	0	0	0	0
1	1	1	1	0	0	0	0	0	0	0	0	0
0	0	0	0	1	1	1	1	1	1	1	1	1
0	0	0	0	0	0	0	0	0	0	0	0	0

Figure 5. 3:Columns 27-39 of target matrix

40	41	42	43	44	45	46	47	48	49	50	51	52
0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0
1	1	1	1	1	1	0	0	0	0	0	0	0
0	0	0	0	0	0	1	1	1	1	1	1	1

Figure 5. 4: Columns 40-52 of target matrix

53	54	55	56	57	58	59	60
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
1	1	1	1	1	1	1	1

Figure 5. 5: Columns 53-60 of target matrix

1	2	3	4	5	6	7	8	9	10	11	12	13
0.9644	0.8933	0.9923	0.9629	0.9810	0.9726	0.9885	0.9833	0.9867	0.9373	0.9469	0.9555	0.9801
0.1642	0.0748	0.0729	0.1404	0.1715	0.0937	0.1855	0.1167	0.1185	0.0840	0.1330	0.1973	0.1220
0.0162	0.0053	2.7402e-04	0.0291	0.0308	0.0202	0.0361	0.0069	0.0021	0.0012	0.0040	0.0210	0.0377
0.0013	0.0096	0.2091	0.0015	0.0013	8.0696e-04	5.3411e-04	0.0029	0.0183	0.0541	0.0237	9.4860e-04	5.5687e-04

Figure 5. 6: training output the first 13 result

14	15	16	17	18	19	20	21	22	23	24	25	26
0.9971	0.9741	0.4663	0.1264	0.0222	0.0253	0.0168	0.1082	0.0601	0.0510	0.0604	0.0141	0.1856
0.0965	0.1336	0.5043	0.8853	0.9440	0.9621	0.9436	0.8993	0.9128	0.9413	0.9329	0.9648	0.9354
0.0021	0.0130	0.0110	0.0446	0.0666	0.0598	0.0915	0.0327	0.0528	0.0685	0.0248	0.0438	0.0170
0.0145	0.0016	0.0062	5.2678e-04	5.1457e-04	5.0818e-04	5.9651e-04	9.9897e-04	4.9227e-04	0.0033	0.0013	8.3388e-04	0.0038

Figure 5. 7: training output the next 13 result

27	28	29	30	31	32	33	34	35	36	37	38	39
0.0597	0.0355	0.0407	0.0814	0.0076	0.0028	0.0060	0.0051	0.0173	0.0021	0.0054	0.0131	0.0068
0.9203	0.9378	0.9470	0.9664	0.0775	0.0549	0.0549	0.0683	0.0676	0.0660	0.0510	0.0689	0.0494
0.0427	0.0563	0.0504	0.0136	0.9209	0.9582	0.9604	0.9712	0.9861	0.9544	0.9633	0.9692	0.9322
9.4708e-04	4.7798e-04	5.5444e-04	0.0036	0.0331	0.0328	0.0446	0.0640	0.0537	0.0324	0.0339	0.0523	0.0917

Figure 5. 8: Columns 27-39 of training output matrix

40	41	42	43	44	45	46	47	48	49	50	51	52
0.0016	0.0015	0.0019	0.0012	0.0034	0.0023	0.0346	0.0085	0.0060	0.0028	0.0135	0.0067	0.0128
0.0747	0.0880	0.0612	0.1144	0.0692	0.0839	0.0368	0.0462	0.0466	0.0563	0.0579	0.0453	0.0482
0.9445	0.9150	0.9570	0.9446	0.9649	0.9185	0.0923	0.1021	0.1398	0.1318	0.2945	0.0758	0.2037
0.0893	0.0815	0.0267	0.0886	0.0423	0.0658	0.9816	0.9704	0.9504	0.9536	0.9468	0.9698	0.9489

Figure 5. 9: Columns 40-52 of training output matrix

52	53	54	55	56	57	58	59	60
0.0128	0.0220	0.0098	0.0315	0.0171	0.0363	0.0284	0.0325	0.0184
0.0482	0.0350	0.0384	0.0372	0.0357	0.0280	0.0316	0.0334	0.0338
0.2037	0.0652	0.0678	0.0894	0.1017	0.1176	0.0841	0.0678	0.7718
0.9489	0.9831	0.9776	0.9823	0.9622	0.9369	0.9698	0.9830	0.1599

Figure 5. 10: Columns 52-60 of training output matrix

1	2	3	4	5	6	7	8	9	10	11	12	13
1	1	1	1	1	0	0	0	0	0	0	0	0
0	0	0	0	0	1	1	1	1	1	0	0	0
0	0	0	0	0	0	0	0	0	0	1	1	1
0	0	0	0	0	0	0	0	0	0	0	0	0

Figure 5. 11: Columns 1-13 of desired output matrix depend on test matrix

14	15	16	17	18	19	20
0	0	0	0	0	0	0
0	0	0	0	0	0	0
1	1	0	0	0	0	0
0	0	1	1	1	1	1

Figure 5. 12: Columns 14-20 of desired output matrix depend on test matrix

1	2	3	4	5	6	7	8	9	10	11	12	13
0.9415	0.9855	0.9863	0.9417	0.9777	0.3790	0.0220	0.0666	0.1111	0.0553	0.0137	0.0034	0.0090
0.4435	0.1559	0.1944	0.2041	0.1627	0.8978	0.9369	0.9499	0.8984	0.9401	0.0797	0.0759	0.0549
0.0012	0.0074	6.4553e-04	0.0159	0.0067	0.0171	0.0101	0.0195	0.0430	0.0286	0.9880	0.9613	0.8402
0.0767	0.0043	0.1102	0.0035	0.0044	0.0014	0.0055	0.0021	7.7432e-04	0.0011	0.0506	0.0561	0.2918

Figure 5. 13: Columns 1-13 of test output matrix

14	15	16	17	18	19	20
0.0054	0.0034	0.0110	0.0189	0.0357	0.0288	0.0230
0.1069	0.0751	0.0372	0.0364	0.0330	0.0360	0.0358
0.9879	0.9805	0.1184	0.0734	0.0818	0.0866	0.0958
0.0704	0.0313	0.9310	0.9783	0.9788	0.9775	0.9736

Figure 5. 14: Columns 14-20 of training output matrix

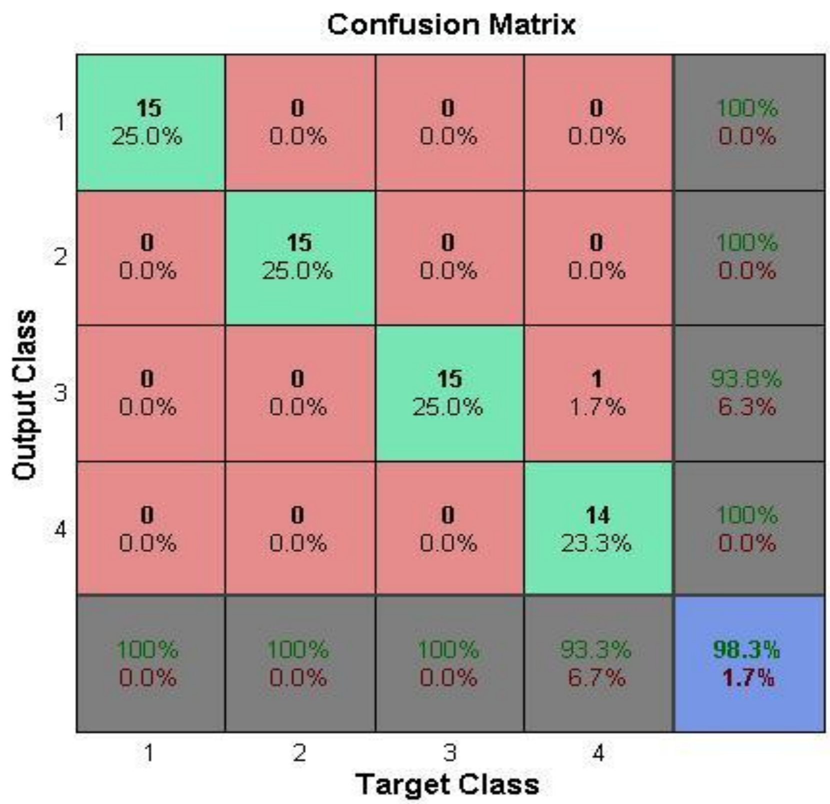


Figure 5. 15: Network confusion matrix

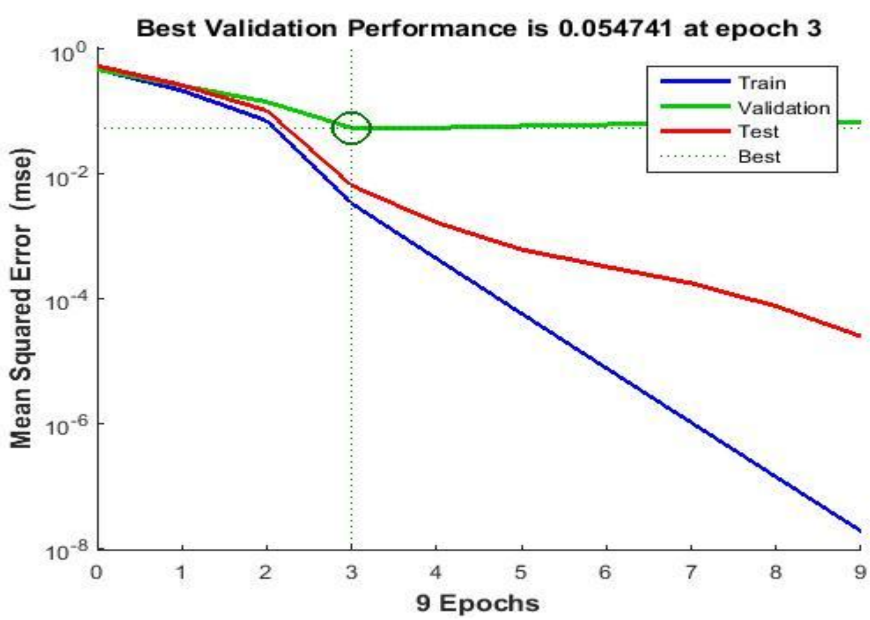


Figure 5. 16: Performance of the neural network

```
function [y] = testNetwork( testSoundFile , myNetwork1)
ResultMatrix = sim(myNetwork1, testSoundFile);
Class1 = (ResultMatrix)
Class1 = round(ResultMatrix)
if Class1(1) == 1
    %for down y=1
     msgbox('the word is down')
    y=1

elseif Class1(2) == 1
    %for left y=2
     msgbox('the word is left')
    y=2
elseif Class1(3) == 1
    %for right y=3
     msgbox('the word is right')
    y=3
elseif Class1(4) == 1
 %for up y=4
  msgbox(' the word is up')
y=4
else
    % GUI ('unable to recognize please retry again')
    msgbox('unable to recognize please retry again')

end
```

Figure 5.16 : testing the network with the word ‘down’

5.2 Discussion of neural network result

Figure 5.7 shows that that maximum error occurred in column 16 which is 0.4 in the first row and 0.5 in the second row. In figure 5.13 and 5.14 the maximum error is 0.4. The confusion matrix in figure shows that the neural network recognized the words by 98.3% of correct classification and 1.7% of incorrect classification. The network failed by 1.7% to recognize the word ‘up’. Figure 5.8 shows that the square of errors is less than 0.1 .These result is satisfy the requirements of the system. After running the code in figure 5.17 with input ‘down’ the message shown in the same figure is displayed.

5.3 Result of UAV simulation

The motion of UAV according to received command is obtained in figure 5.185.19,5.20 , 5.21 and 5.22

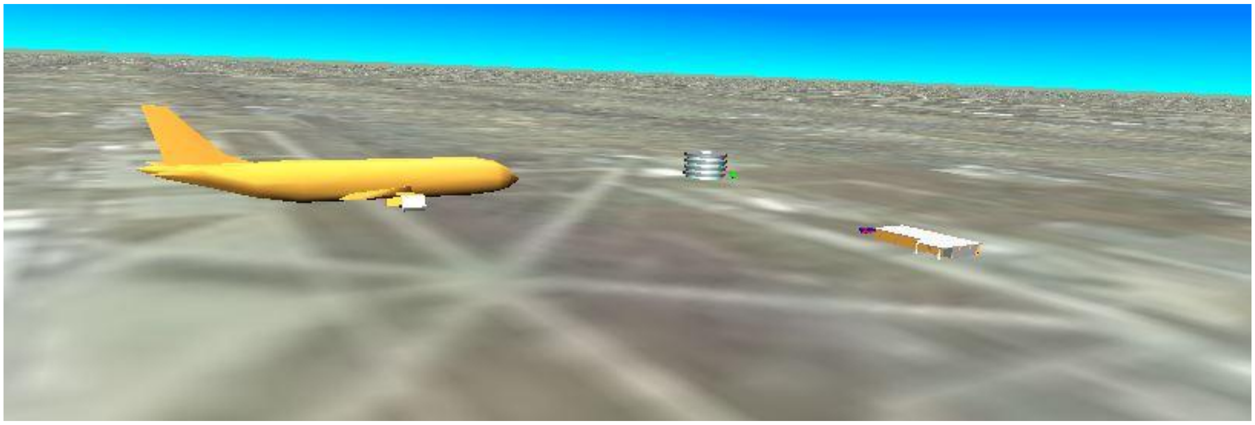


Figure 5. 17: The normal attitude of UAV

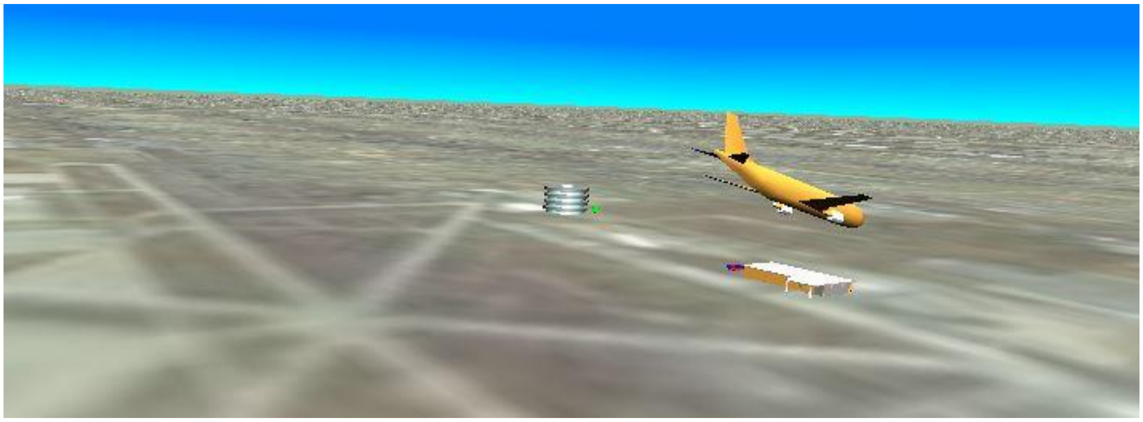


Figure 5. 18: UAV moving downward

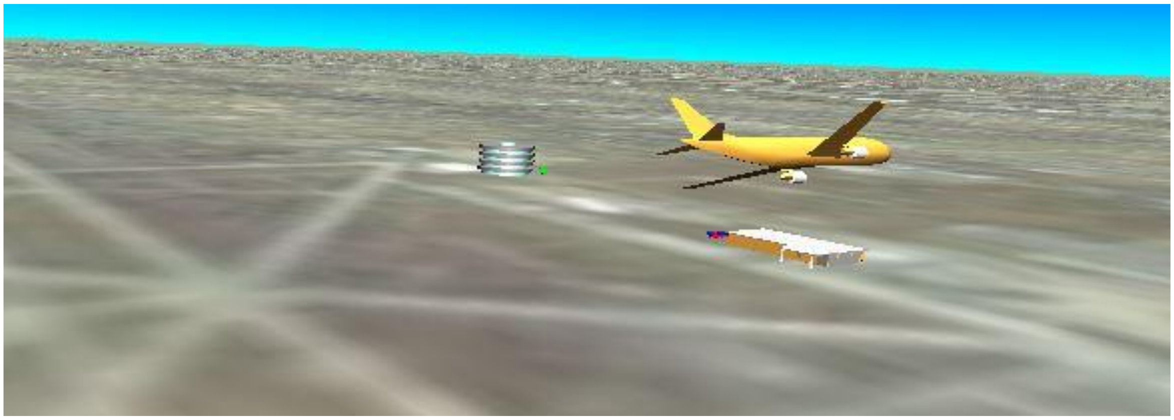


Figure 5. 19: UAV moving to the left

Figure 5. 20: UAV moving to the right

Figure 5. 21: UAV moving upward

5.4 Discussion of UAV simulation

It is noted from figures 5.19, 5.20, 5.21, 5.22 that the motion of the UAV is the, down, left, right, and up respectively when these commands are received in the same order.

5.5 Result of designing interfaces

Figures 5.23, 5.24, 5.25, 5.26, show the interfaces of the system

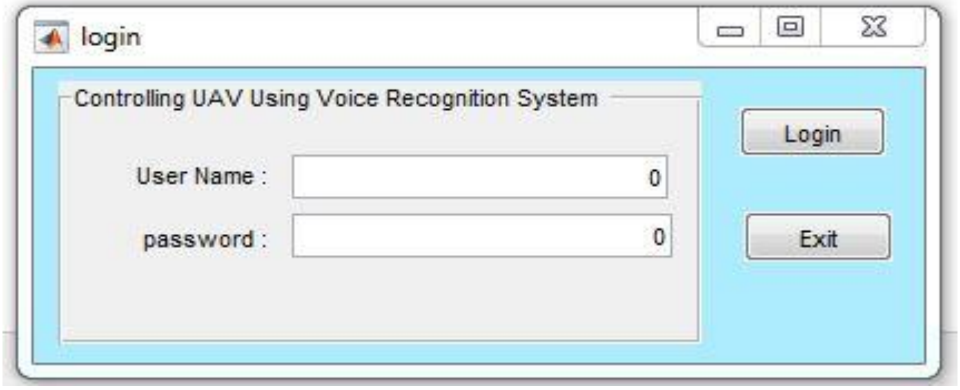


Figure 5. 22: login interface

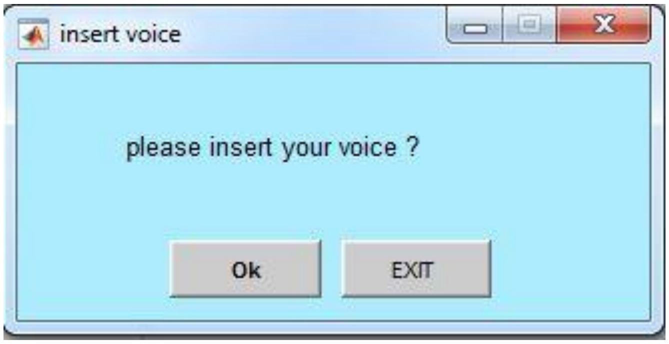


Figure 5. 23: Insert the voice interface

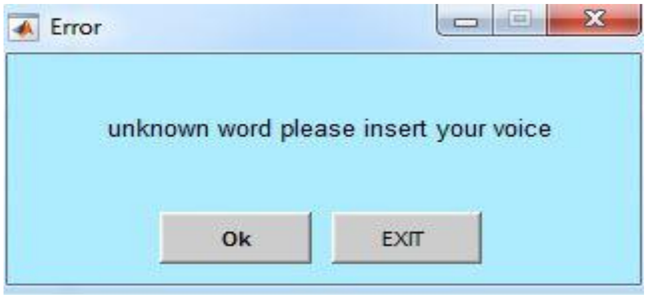


Figure 5. 24: Error message due to unknown word interface

Figure 5.26: Error message due to unknown word interface

CHAPTER SIX

CHAPTER SIX

CONCLUSION AND RECOMMENDATIONS

6.1 Conclusion

In this thesis speech recognition is proposed to control UAV by using neural network to recognize commands and the microcontroller code was written to convert these commands to actual motion. A code and simulation on the
Matlab Simulink environment was undertaken. The good result of recognize word is and the expected UAV motion is obtained. Also the interfaces of this system are designed.

6.2 Recommendations

We recommended to improve this work with the following points:

1. Applying this techniques on multiple UAVs.
2. The simulation has been applied in this research, it is better to test on the real world
3. Neural network is implemented in this work. It is better to test other techniques to improve the performance of all the required
parameter.

Reference

[1] Alaa Hassan Mahmoud and Salma Alzaki Ali, Speech To Text Conversion, Sudan University of Science &Technology Faculty of Computer Scinese & Infromation Technology, 2014.

[2] Artificial_Intelligence-A_Guide_to_Intelligent_Systems.pdf

[3] download-pdf-ebooks.org-ku-11303 –pdf.

[4]cmsdu.org/documents/Microcontroller_Microprocessor_Differences.pdf

[5] www.coursehero.com/file/18407294/labreportweek 10/ 10/10/2017

[6] https://www.quora.com/What-is-the-difference-between-a-microprocessor-and-microcontroller 17/10/2017 11:35 AM

[7]Seemaramdhaf and shilpajoshi, robot controlled by android device through the speech recognition,PES College of Engineering,Pune, India,3 rd. April, 2016.

[8]safaa omer mohammed,Voice recognition by using machine learning a case study of some rules،

،Sudan University Of Science And Technology College Of Graduate Studies .

[9]maye abdalslam and hiba abdall rahman,PC Control OF RC Car,Sudan University Of Science And Technology

College Of Computer Science and Information Technology, 2011.

[10] Abu-Bakr and Osama and Mortda,Control the wireless plane by computer,Sudan University Of Science And Technology CollegeOf Computer Science And Information Technology,2011.

[11] Ashraf Mohammed, SOUND RECOGNITION FOR SELECTED VERSES OF THE HOLY QURAN ,Sudan University Of Science And Technology College Of Graduate Studies ,May/2012.

[12] Alexander Murphy, Implementing Speech Recognition with Artificial Neural Networks, Algoma University Sault Ste. Marie, Ontario April 11, 2014.