MEF UNIVERSITY
FACULTY OF ENGINEERING

DEPARTMENT OF ELECTRICAL AND ELECTRONICS ENGINEERING

SMART HOME ASSISTANT: JARVIS

Senior Design Project

Utku Mutlu (041502031)

Alper Salik (041502005)

Hasan Pektaş (041502014)

Advisors: Assoc. Prof. Ebru Arısoy Saraçlar

Assoc. Prof. Tuba Ayhan

2020

MEF UNIVERSITY
FACULTY OF ENGINEERING

DEPARTMENT OF ELECTRICAL AND ELECTRONICS ENGINEERING

Project Title   : Smart Home Assistant: Jarvis
Student's Name   : Alper Salik, Hasan Pektaş, Utku Mutlu
Date   : 08/06/2020

I hereby state that the design project prepared by Alper Salik, Hasan Pektaş and Utku Mutlu has been completed under my supervision. I accept this work as a "Senior Design Project".

08/06/2020
Dr. Ebru Arısoy Saraçlar

Dr. Tuba Ayhan

I hereby state that I have examined this senior design project by Your Name (Title Format). This work is acceptable as a "Senior Design Project".

08/06/2020
Department Chair's Name

Head of the Department of
Electrical and Electronics Engineering

ACADEMIC HONESTY PLEDGE

In keeping with MEF University Student Code of Conduct, I pledge that this work is my own and that I have not received inappropriate assistance in its preparation. I further declare that all resources are explicitly cited.

NAME	DATE	SIGNATURE
Utku Mutlu	08/06/2020	

NAME	DATE	SIGNATURE
Alper Salik	08/06/2020	

NAME	DATE	SIGNATURE
Hasan Pektaş	08/06/2020	

ABSTRACT

SMART HOME ASSISTANT: JARVIS

Utku Mutlu
Alper Salik
Hasan Pektaş

MEF UNIVERSITY
Faculty of Engineering
Department of Electrical and Electronics Engineering

Advisors: Ebru Arısoy Saraçlar
             Tuba Ayhan

JUNE, 2020

Smart home systems integrate many new technologies to improve people's quality of life through the home network. In this project, a smart home assistant whose name is Jarvis is developed like Google Home or Amazon' Alexa. The assistant presented for people with disabilities. This smart home assistant understands previously defined Turkish commands, and then it performs the tasks associated with the given commands. The system has two main parts which are Automatic Speech Recognition (ASR) and performing tasks related with commands. The Mel-Frequency Cepstrum Coefficient (MFCC) and Dynamic Time Warping (DTW) methods were used for the ASR part. In the second part two circuits were designed. One of the circuits works with 2 different switches that are set up to turn the lamp on and off because when the user turns off or on the lamp the user can on or off it from the normal switch. The second circuit has an IR Receiver to decode signals of Tv remote controller and an IR Transmitter to send this signal to Tv by speech command. A package for hotword detection was used. The system is always listening to users and when users say 'Jarvis' command, it gets active and waits for next commands. The accuracy of the system was tested. The results showed that accuracy of performing commands of design is more than 80%.

Keywords: Mel Frequency Cepstrum Coefficient, Dynamic Time Warping, smart home, speech recognition

ÖZET

AKILLI EV ASİSTANI: JARVİS

Utku Mutlu
Alper Salik
Hasan Pektaş

MEF ÜNİVERSİTESİ
Mühendislik Fakültesi
Elektrik-Elektronik Mühendisliği Bölümü

Tez Danışmanı: Ebru Arısoy Saraçlar
Tuba Ayhan

HAZİRAN, 2020

Akıllı ev sistemleri, insanların yaşam kalitesini ev ağı üzerinden iyileştirmek için birçok yeni teknolojiyi entegre eder. Bu projede, adı Jarvis olan bir akıllı ev asistanı, Google Home veya Amazon'un Alexa gibi, geliştirildi. Bu asistan engelli insanlar için sunuldu. Bu akıllı ev asistanı daha önce tanımlanmış Türkçe komutları anlar ve daha sonra verilen komutlarla ilişkili görevleri yerine getirir. Sistem, otomatik konuşma tanıma ve komutlarla ilgili görevleri yerine getiren iki ana kısma sahiptir. Otomatik konuşma tanıma bölümü için mel frekanslı cepstrum katsayısı ve dinamik zaman çözgü yöntemleri kullanıldı. İkinci bölümde iki devre tasarlandı. Devrelerden biri, lambayı açmak ve kapatmak için ayarlanmış 2 farklı anahtarla çalışır, çünkü kullanıcı lambayı kapattığında veya açtığında kullanıcı normal anahtardan açabilir veya kapatabilir. İkinci devre, TV uzaktan kumandasının sinyallerini çözmek için bir IR alıcısına ve bu sinyali konuşma komutuyla TV'ye göndermek için bir IR vericisine sahiptir. Hotword algılama için bir paket kullanıldı. Sistem her zaman kullanıcıları dinliyor ve kullanıcılar 'Jarvis' komutunu söylediğinde aktif hale geliyor ve bir sonraki komutları bekliyor. Sistemin doğruluğu test edildi. Sonuçlar, tasarım komutlarının gerçekleştirilmesinin doğruluğunun %80'den fazla olduğunu gösterdi.

Anahtar kelimeler: Mel frekans kepstrum katsayısı, dinamik zaman çözgü, akıllı ev, konuşma tanıma

TABLE OF CONTENTS

LIST OF TABLES

1

LIST OF FIGURES

LIST OF ABBREVIATIONS

DTW   Dynamic Time Warping

MFCC   Mel Frequency Cepstral Cepstrum

HMM   Hidden Markov Model

IR   InfraRed

LIRC   Linux InfraRed Control

1. INTRODUCTION

Motivation, aim ,target and impact of our smart assistant were included in ıntroduction.

1.1. Motivation

The Smart Home Assistant is a connecting device of your home, such as lighting, cameras, sensors and alarms, to work together. This device helps to create a network to control devices through Wi-Fi, Bluetooth and other connections. Managing all your devices from one place is tremendous comfort.

Smart home systems [1] are technological developments that makes it significantly easier in our lives. Smart home systems could offer people simplify ordinary tasks for a comfortable lifestyle. The main purpose of the systems in your home is to conduct the automation system, the time-consuming processes and to perform the operations that the user can not perform at that moment. When the people come home exhausted after a long hard working day, smart sockets help them switch their lights on or off, etc. by their voice commands. In addition, it was particularly beneficial to people with disabilities, creating a lifestyle that was previously impossible.

1.2. Broad Impact

1.2.1. Public Health, Safety, and Welfare

It is always a priority in lives to keep family and property safe. Smart homes are the most widely used systems in the area, alerting you to problems before disasters happen.

With smart home systems, the security level in the homes can be strengthened. With the increase in the number of sensors, it can detect security-violating situations and not provide unnecessary alarms, even monitoring suspicious situations. It is more reassuring than the next generation of security systems.

The security is not in terms of burglaries. When a smart house turns off the power from the iron user left in the socket, the user doesn't have to think about whether the user left the iron on after you left the house. Moreover air quality sensors monitor various factors in the air to help users to keep healthy. They might warn users of unhealthy conditions or turn on a ventilation system.

1.2.2. Global, Cultural, Social, Environmental, and Economic Impact

Light sensors are essential parts of smart homes. With this technology, the users can enable the lights to turn on automatically when the room light is dark and turn off automatically when the room light is bright. Smart home technologies offer advantages in lighting as well as heat and energy. These various advantages are keeping the house temperature at the desired level, saving energy by lowering the temperature while the users are sleeping or when users are not at home, or heating each room to a separate degree. Therefore these systems have a positive impact on the environment by reducing the energy consumption.

2. PROJECT DEFINITION

In this project a smart home assistant like Google Home or Amazon's Alexa was designed. The name of the assistant is Jarvis. Jarvis understands previously defined speech commands. Then it performs the tasks related with the given commands. Designed to control some home appliances using user-provided command templates lights on/off, television on/off. Jarvis responds to certain commands by playing audio files which are previously recorded. When the user says 'Jarvis', it responds to the user by saying ' Yes Sir'.

The aim of this project is to build a low-cost personal assistant to make life easier for users in their homes. Furthermore the personal assistant helps users to perform some work for people with disabilities. The target in this project was software and hardware systems. This software system understands the users speech commands and responds to commands and this hardware system performs the tasks involved with given commands..

This project includes two steps. In the first step, an Automatic Speech Recognition [2] system was designed to compare previously recorded speech commands of users and real-time commands of users. In the second step, performing the tasks step was designed associated with given commands.

In the first step Mel Frequency Cepstrum (MFCC) [3] and Dynamic Time Warping (DTW) [4] methods were used to generate speech signals. A few sounds which were 16-bit and 16 kHz at first were recorded. Then the MFCC command in Python to find Mel coefficients of audio signals was used. After the MFCC process, the DTW part was applied to

find the distance of the Mel coefficient vectors of template speech commands and Mel coefficient vectors of real time speech commands to see if the sounds match.

In the second part a circuit to turn lıght off and lıght on that works with 2 different switches was set up. Because when the user switches off or on the lamp via the smart home system, the user can switch it off or on from the normal switch. A remote controller designed to communicate with the television to control on, off, volume up, volume down and channel switching functions of television. The IR transmitter and receiver used that part.

At last a word detection system added to design. We used a hotword package called Snowboy [5].This package trains our command templates. This system uses voice activity detection and deep learning technologies. The hotword is 'Jarvis' command for this design. Our system is always running in listening mode and when users say Jarvis, it wakes up. This hotword detection system works offline.

In the first design processing time of our design was about 2 minutes. In the second design the DTW code was changed to reduce process time. Normally to find the best path value all the DTW values were holding. However, this is required only in the" Jarvis " command, which is required for the system to wake up. So we decided not to keep all DTW lengths. 80% speed up in the system was achieved when compared with previous design. The language of commands was English in first design. The language of commands changed to Turkish. Because our first target user profile is people who know Turkish. Moreover Turkish is easier to process and train than English.

3. THEORETICAL BACKGROUND

Smart home system project includes software and hardware sections. The software part includes command recognition and the hardware part is able to perform the tasks given by commands. In this part of the project we focus on command recognition.

3.1. Literature Survey

3.1.1. Speech Recognition

Speech recognition is an interdisciplinary subfield of computational linguistics that develops methodologies and technologies that computer software program or hardware device with the ability to decode the human voice. Speech recognition is commonly used to operate a device, perform commands, or write without having to use a keyboard, mouse, or press any buttons. Raj Reddy was the first person to take on continuous speech recognition as a graduate student at Stanford University in the late 1960s [6]. Previous systems required users to pause after each word. Reddy's system issued verbal commands to play chess. During this time, Soviet researchers invented the dynamic time warping (DTW) algorithm and used it to create a recognizer that could run on a 200 word vocabulary. DTW processed the speech by dividing it into short frames and processing each frame as a single unit. Achieving speaker independence remained unresolved during this period. Raj Reddy's students James Baker and Janet Baker began using the Hidden Markov model (HMM) for Speech Recognition ın the 1970s [6]. The use of hmms allowed researchers to combine different sources of information, such as acoustics, language, and syntax, in a unified probabilistic model.

DTW is a method that calculates an optimal match between two given sequences. The basis of DTW is found on the computations of distance or confusion matrix between two-time series. The similarity or dissimilarity of two sequences is typically calculated by converting the data into vectors and calculating the Euclidean distance between those points in vector space.

Hidden Markov models [7] are statistical models that are closely related, as the name already suggests, to Markov models. In contrast to Markov models, where the states are directly visible, the states are not directly visible in the case of HMM. In other words, each hidden state holds different probability distributions that produce observable output. When applying the HMM, a multivariate Gaussian emission distribution in each hidden state is

assumed. This is essentially since there are efficient parameter estimators for this special case. This distribution assumption, however, does not hold true for many data sets HMM can be viewed as an extension of Markov chains. The only difference compared to common Markov chains is, that the state sequence corresponding to an observation sequence is not observable but hidden. In other words, observation is a probabilistic function of the state, and the underlying sequence of states itself is a latent stochastic process. This means that the sequence of states can only be observed indirectly through another stochastic process that emits an observable output. The mean of stochastic is assumed depending on the current state to future states.

3.1.2. Smart Home Systems

Smart home technology is connecting devices via a network, most commonly a local LAN or the internet and using devices from one center. The technology was originally developed by IBM and first contemporary smart home technology products became available to consumers between 1998 and the early 2000s [8]. Google Home and Amazon Echo are the largest manufacturers of smart home systems on the market. These systems use speech recognition technology over the internet connection. In the Google Home system the devices enable users to speak voice commands to interact with services through Google Assistant, the company's virtual assistant. Users can listen to music, control playback of videos or photos, or receive news updates entirely by voice.

The trigger word detection system is used widely for smart home systems. The systems which wake up on listening to a specific word are called trigger word detection systems. Active words are the words on which we want to wake up our system which are called trigger words. Passive words are those words which should not wake up our system. All other words other than our trigger word and background noises will act as negative words.

3.2. Solution Methods

One of the most important problems in speech recognition is that different people say the same word differently, as well as the same person doesn't say the same word in the same way at different times. Even when the same user repeats the same word, it may not resemble previous vocalizations. So the speech recognition was designed to solve this problem. After searching several articles, controlling the system with Raspberry Pi was decided. Then we decided to use the DTW algorithm but first they need to use MFCC for feature extraction. After getting MFCC, DTW algorithm compares the commands that we have already defined in the system and finds the minimum distance and chooses the command. Five different commands from five different people were recorded and defined on the computer. These commands were used to compare with DTW. The algorithm was written with Python language because it is the most useful and easy language.

In first design our processing time of the ASR part was about 2 minutes. Our system held all of DTW values to find the best path value. We found that this method was only necessary for the "Jarvis" command to wake up the system and we kept only the DTW lengths of the Jarvis command. After this process, the system's decision-making process accelerated.

4. DESIGN

Block diagram of the software part of our project is shown in Figure 1. This block diagram of ASR which includes Feature Extraction, Dynamic Time Warping and Decision modules is shown in Figure 2. A block diagram of Execution of Commands includes Coding on Raspberry PI, Designing Circuits and Implementation of Circuits shown in Figure 3. We used the MFCC method to remove acoustic properties in the speech signal. Then we used the DTW method to find the similarity between vectors of recorded templates and vectors of real time speech signal. In the decision part check the system does correctly recognize the command.
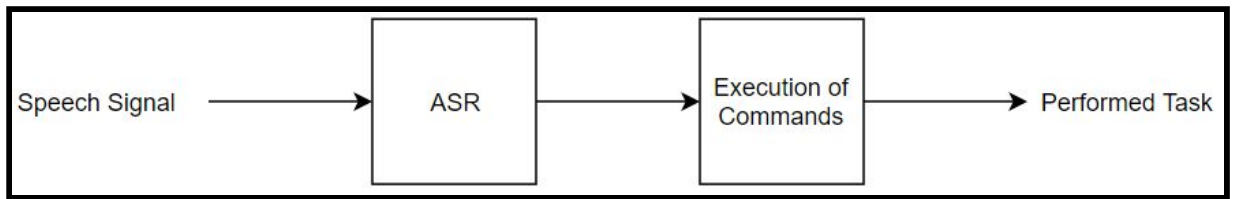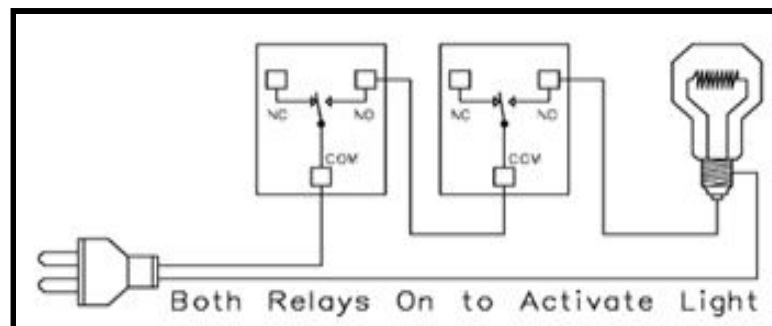
Figure 1. Block Diagram of Smart Home Project.
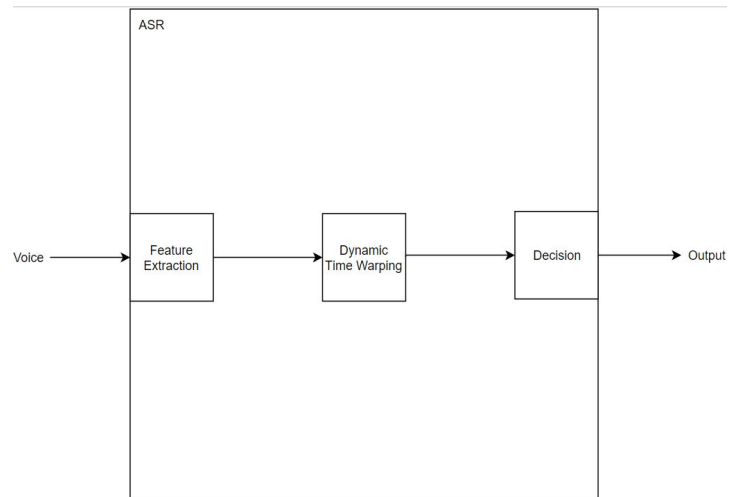
Figure 2. Block Diagram of Automatic Speech Recognition(ASR) Part.

Figure 3. Execution of Commands Circuit

4.1 System Components

4.1.1 Data

We collected data to check the accuracy of the system and to add a template that can be checked into the system. All of our audio files had been in wav format, 16kHz and 16 bit PCM. After receiving the data, we had extracted MFCC features. After the mfcc process, we had had 13 frames. We had compared these frames with the template by doing dynamic time warping. First we had compared the commands that the same person said at different times with the template. In contrast, we had changed the speed and added noise to some sounds. When we had examined the results of this situation in excel, we had found that accuracy is high. There are template files on the x axis and voice recordings of the same person at different times on the y axis. These parts were done on EE 491.

More datas was collected to find a more efficient rate. Apart from these data, the sounds to output the Jarvis sound were recorded. According to the command these voices were given, we are told back as output.To get the Jarvis sound, the sound was divided into frames and inverted the frames. In this way, a robotic sound was made. All of the audio files had been in wav format, 16kHz and 16 bit PCM too. These parts were done on EE 492.

4.1.2 Feature Extraction

The removal of acoustic features in the audio signal is called Feature Extraction. Mel-Frequency Cepstrum Coefficient (MFCC) [3] is one of the most widely used Feature Extraction methods in speech recognition.

The MFCC process consists of the following steps and block diagram of MFCC shown in Figure 4.

1. Frame Blocking:

The sound signal characteristics remain stable over small time intervals. Therefore, sound signals are processed in short time intervals. The signals are divided into frames, lengths of these frames about 50ms.

2. Windowing:

The purpose of windowing is to eliminate discontinuities at the beginning and end of the frame. One of the most commonly used windowing functions at this stage is the Hamming function.

3. Fast Fourier Transform (FFT):

In each frame the Fast Fourier Transform is taken and this process passes the set of definitions from time space to frequency space.

4. Mel-Frequency Warping:

Mel frequency scale is a scale that shows the perception of the change in the sound frequencies of the human ear. While the perception of sounds up to 1000 Hz is linear, the perception of change becomes logarithmic as the frequency increases. This stage of speech recognition is used as a band pass filter.

5. Cepstrum:

As the final stage of Feature Extraction, each frame is inverted fourier transformed and returned from frequency space to time space again.As a result of this process, Mel-Frequency Cepstral Coefficients (MFCC) are obtained.

In our design the waveform of the first audio file is extracted. In the next step, the audio file is defined as a vector. Values defined as vectors are divided into frames of 13 values. The next frame is taken starting from half of the values in the previous frame. This is to prevent the loss of sound values. FFT (Fast Fourier Transform) is applied to the values allocated to frames. The aim of FFT application is to measure the frequencies present in the sound. Once these frequencies become measurable, they are passed through the filterbank. Thus, each signal carries its own value. Finally, the discrete cosine transform (DCT) is applied to obtain MFCC frames.
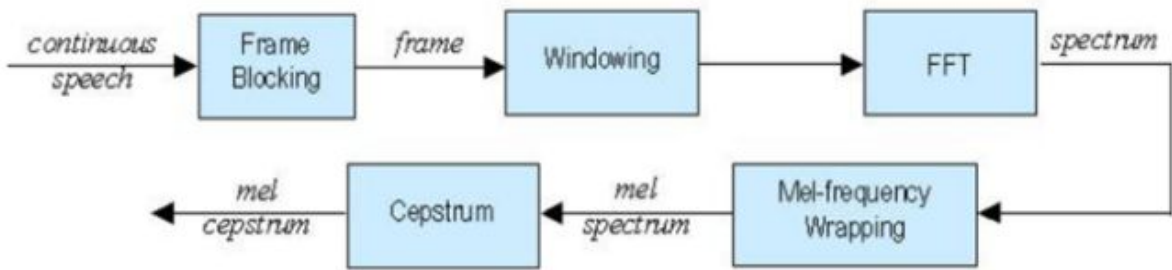
Figure 4. Block diagram of MFCC

4.1.3 Dynamic Time Warping (DTW)

The Dynamic Time Warping [4] method shown in Figure 5 helps us to find the similarity of two speech commands to prevent this problem. In this method by matching the values of the vectors closest to each other, the sum of the distances is kept at the smallest value. With this method, similarity measurement can be made for vectors of different dimensions.

For DTW method firstly Euclidean distances are calculated. The Euclidean distance between two values in a dimension is found by taking the square root of the square of the difference of these two values, it is actually the absolute value of the difference. The cumulative distance is the sum of the distance of a value to its mapping, and the distance of all the values that match before its mapping. When calculating the cumulative distance, the smallest of the cumulative distances calculated for pairs of a unit vector behind these elements is added to the calculated Euclidean distance between each pair of vectors. Thus, when the last matching elements are reached, the last cumulative distance value calculated is the smallest possible value. In our design we have 13 vectors in each frame of our speech signals. We calculated the smallest possible value using the DTW method. Our design compares the distance value of the users real time speech signals and the distance value of our template signals to understand correct command.

The example to find minimum value with DTW and its path shown in Figure 5. According to Figure 5, T values are vectors of recorded sound and C values are vectors of real time sound. Moreover I have vector numbers of C and J are vector numbers of T.
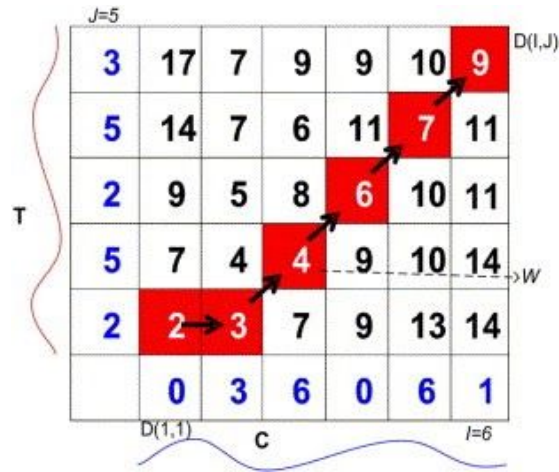

Figure 5. Example of DTW

In Figure 5, to compute D (3,4), firstly we find the distance of 5 and 6 by 6-5 = 1. Then we have to add the previous path to the value of 1. For this, we have to choose the minimum of D (3,3), D (2,3) and D (2,4). This value is 5 at point D (2,3). Then we add 5 to 1 and reach 6.

4.1.4 Decision

While doing the decision, we looked at the dynamic time warping scores, and these small scores indicate that the input speech and the template speech are the same. In response to the command, we receive feedback that states that the system understands the command.On EE 492 we established a communication network with the devices over Raspberry Pi to fully execute the command. The system is always in sleeping mode until the user says "Jarvis", then the system wakes up and waits for another command.

4.1.5 Circuit Designs
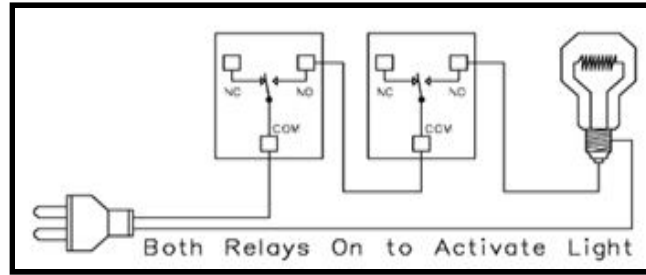
Figure 6. Block Diagram of 2 switch lamp with relay

We set up a circuit [9] shown in Figure 6 that works with 2 different switches to turn the lamp on and off because when we turn off the lamp the user can open it from the normal switch. Our first switch is a regular lamp switch and the other one is able to turn the lamp on and off with the signal from raspberry pi with the command on or off the lamp. To operate this we used 2 channel relays. This relay works with 5V. When we give 0V output to the relay input pin it opens the switch. We have 2 input pins in the relay, one of them takes a signal from a raspberry pi pin and the other one takes a signal from a regular switch. When we say "lambayı aç" raspberry pin 16 gives 0V and when we say "lambayı kapat" raspberry pin 16 gives 3.3V.
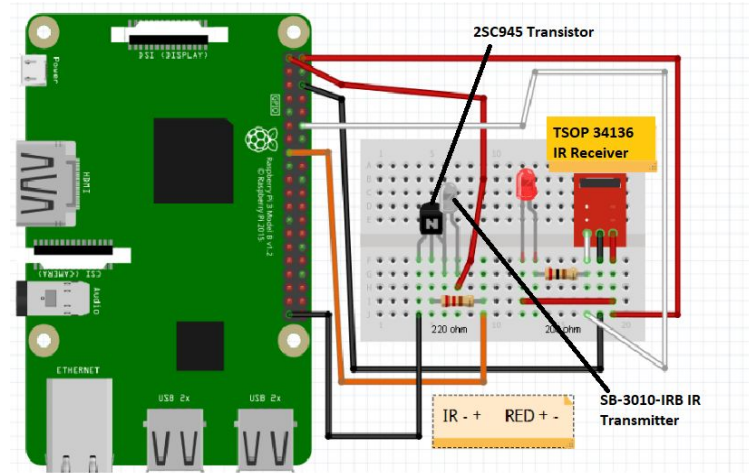
Figure 7. Block Diagram of IR Receiver-Transmitter circuit [10]

We need an IR Receiver to decode Tv remote controller signals and we need an IR Transmitter to send this signal to Tv by speech command. We used TSOP 34136 IR Receiver and SB-3010-IRB IR Transmitter. Also we used a 2SC945 Transistor to amplify signals. Also we put a led to see that the signals are sent and received [10].

5. IMPLEMENTATION AND TESTING

5.1. Implementation

5.1.1 Software

The ASR part was firstly implemented on a computer running Windows 10 with Python 2.7. We used a lot of Python libraries [11] for MFCC and DTW methods. Template mfcc values of audio recordings kept in txt files. So every time the program runs, it doesn't do a lot of mfcc operations. This reduces the slowdown of the program. The software part of this project moved to Raspberry Pi 3. When we started working on Raspberry Pi 3, we found that the program slowed down a lot. The reason for this was that the Raspberry Pi 3 operates at 1.4 GHz. Since this value is 3.2 GHz on our Windows 10 computer, the program was running 3 times slower. In order to solve this problem, we made some changes to our DTW application.

Normally when applying DTW, we were holding all the values to find the best path value. Then we kept only the DTW values of 'Jarvis' command. Speed increase shown in the system.

We used the Snowboy API [12] for hotword detection. Snowboy is an embedded and real-time, always-listening but off-line, and highly customizable hotword detection engine that runs on Raspberry Pi, (Ubuntu) Linux, and Mac OS X. We chose the word Jarvis as the trigger word. We trained the system with the sounds in which the word Jarvis was said. We have added this system to our code. When we say Jarvis, our system wakes up and responds to us. Then it starts listening to us for the command he will perform and then performs this command.

We had to send an infrared signal to control the remote. We did this with the lirc package in linux. LIRC [13] is an open source package that allows users to receive and send infrared signals with a Linux-based computer system.With LIRC and an IR receiver the user can control their computer with almost any infrared remote control. The user may for instance control DVD or music playback with their remote control.

We used the PIN code to turn the lamp on and off. If the lamp opening process will be done, we send the required pin (gpio) 0V (low) so that the led is turned off. If the led shutdown process is going to be done, we send the required pin (gpio) 5V (high) so that the led is turned on. According to the voltage value sent, the relay was triggered so that we brought the led to the state we wanted.

Actually, pyaudio, wave, mfcc, sys, numpy, pyttsx3, webrtcvad, collections, pyaudio, python_speech_features, GPIO, mfcc, array, os, time, scipy.io.wavfile as wav and euclidean libraries were used.

The GPIO library for using general purpose input-output pins.

The snowboy library is used to hotword detection a wav file.

The Pyaudio library is used to record a wav file.

The Wave library is used to open and edit a wav file. Gets files as input and files came out as a signal.

The mfcc function takes the wav file as input and the sampling rate of that file. It divides this file into frames, and outputs mfcc after output. There are 13 vectors per frame.

The numpy library is a library to facilitate array operations in python. It assigns the values it receives as input to the array and edits these values. Create 2-dimensional arrays.

The sys library provides access to some variables used or maintained by the interpreter and to functions that interact strongly with the interpreter.

The pyttsx3 library is a text-to-speech conversion library in Python.

The webrtcvad library is a python interface to the WebRTC Voice Activity Detector

The collections library implements specialized container data types providing alternatives to Python's general purpose built-in containers.

The time library provides various time-related functions. For related functionality.

The scipy.io.wavfile returns the sample rate (in samples/sec) and data from a WAV file

The euclidean library computes the Euclidean distance between two 1-D arrays.

The os library helps to use the terminal. The command to terminal from python was given.

5.1.2 Hardware

In the hardware part of this project Vcom M727 USB Microphone, Raspberry Pi 3 B+, Mini USB Speaker, 5V 2-channel relay, 83.5mm x 54.5mm x 8.5mm Breadboard, SB-3010-IRB IR Transmitter, TSOP 34136 IR Receiver and 2SC945 Transistor used.

- Vcom M727 USB Microphone: The Vcom M727 USB Microphone shown in Figure 8 was used to high quality recording templates and testing the real time voices with templates.

Figure 8. The Vcom M727 USB Microphone

- Raspberry Pi 3 B+: Raspberry Pi 3 B+ [14] is shown in Figure 9, a small-sized computer that carries the operating system on the card in micro SD. Uses a customized operating system called Raspbian. Python programming language is used. Its processor is 1.4 GHz, 4-core 64-bit.

Figure 9. The Raspberry Pi 3 B+ [14]

- Mini USB Speaker: The Mini USB Speaker shown in Figure 10 was used to hear the response of the system.

Output: RMS 3W+3W

SIN Ratio: 65dB Isolated

Rate: 45dB

Impedance: 4 ohms

Power: USB 5V

Speaker Unit: 2.5Wx2

Figure 10. The Mini USB Speaker

- 5V 2 channel relay: A relay is defined as an electrically operated switch; they control circuits by a low-power signal or when several circuits must be controlled by one signal. The 5V 2 channel relay shown in Figure 11[15].

- ■ **Input:** Vcc, connected to the 5V current on the Raspberry Pi Board, GND, connected to the ground and 2 digital inputs. (In1 & In2)
- ■ **Output**: The 2 channel relay module could be considered like a series switches: 2 normally Open (NO), 2 normally closed (NC) and 2 common Pins (COM)
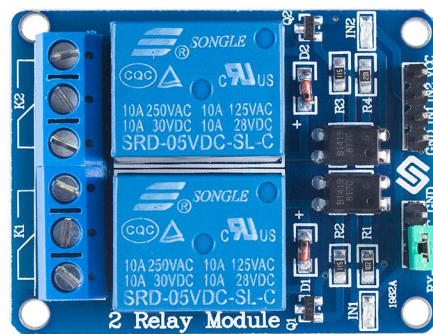
Figure 11. The 5V 2 channel relay [15]

- 83.5mm x 54.5mm x 8.5mm Breadboard: The Breadboard is the tool on which we design and test our circuits. The interior design of the Breadboard consists of metal clamps that are positioned vertically and horizontally connected to each other.

- SB-3010-IRB IR Transmitter: The SB-3010-IRB IR Transmitter [16] shown in Figure 12 was used to send infra-red rays to enable communication from Raspberry Pi to receiving devices. It sends the commands for the task it wants to perform to the receiver.

Figure 12. The SB-3010-IRB IR Transmitter [16]

- TSOP 34136 IR Receiver: The TSOP 34136 IR Receiver [17] shown in Figure 13 was used to receive infra-red rays to enable communication from receiving devices to Raspberry Pi. After it receives the command from the transmitter, it responds to feedback.

- Details Carrier Frequency: 36 kHz
- Transmission Distance: 45 m
- Output Current: 5 mA
- Operating Supply Voltage: 2.5 V to 5.5 V

Figure 13. The TSOP 34136 IR Receiver [17]

- 2SC945 Transistor: The C945 is a Bipolar audio frequency NPN transistor [18] shown in Figure 14. It was used for pre-strengthening purposes of the remote controller circuit. It has a good gain value of a maximum of 700 and is highly linear.

- Continuous Collector current (IC) is 150mA

- Collector-Emitter voltage (VCEO) is 50 V

- Collector-Base voltage (VCB0) is 60V

- Emitter Base Voltage (VBE0) is 5V

- Transition Frequency is 150MHz
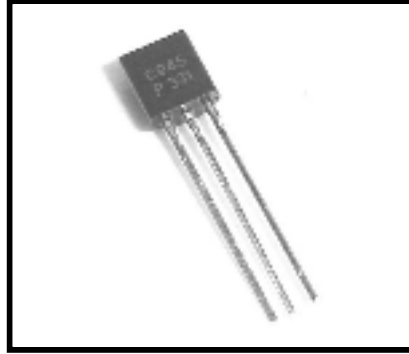
Figure 14. The C945 NPN transistor [18]

5.2. Testing

The project's work was tested with 10 different sounds. At first the voices of the same people to each other were tested. In this comparison 82% accuracy was achieved on EE491. The application's work with 10 more different sounds was tested. In this comparison 80% accuracy was achieved on EE492. Because we changed the commands English to Turkish. Also we used fastdtw library for computing DTW. This library computes the minimum distance without holding all datas.Accuracy of performance of the system by given commands are shown in Table 1. Accuracy of receiving signals for TV commands are shown in Table 2.

According to these results, DTW provides a normal level of accuracy for voice recognition. In noisy environments or when we replace the microphone, these rates decrease. Our system also had processed commands for up to 2 minutes. But after changing DTW, our process time decreased to 10 seconds.

Table 1. Accuracy of performance of system by given commands

COMMANDS	ACCURACY
Jarvis	90%
Televizyonu aç	75%
Televizyonu kapat	70%
Kanal yukarı	75%
Kanal aşağı	80%
Sesi aç	80%
Sesi kapat	75%
Nasılsın	90%
Lambayı aç	80%
Lambayı kapat	80%

Table 2. Accuracy of receiving signal for TV commands

COMMANDS	ACCURACY
Televizyonu aç	65%
Televizyonu kapat	55%
Kanal aşağı	60%
Kanal yukarı	50%
Sesi aç	75%
Sesi kapat	65%

Apart from that, we can command the television by voice control. We do a voice command to the controller we use to do this, and the controller generates the appropriate signal according to this voice command and sends it to the receiver. The average accuracy of the commands are 80%. Also the accuracy of receiving signals for TV commands are 60%. In addition, the written code cannot be optimized with any other control than the one produced, which is one of the challenges we face. The tests were created by taking 20 attempts from each command.

5.3. Cost Analysis

- Vcom M727 USB Microphone 24 TL
- Raspberry Pi 3 B+ 270 TL
- Mini USB Speaker 30 TL
- 5V 2 channel relay 8.57 TL
- 83.5mm x 54.5mm x 8.5mm Breadboard 10 TL
- TSOP 34136 IR Receiver 5.88 TL
- SB-3010-IRB IR Transmitter 0,56 TL
- 2SC945 Transistor 0,27 TL
- Total cost 349,28 TL

6. RESULTS

6.1. Software

Python is a convenient programming language for implementing the proposed system because it has a lot of libraries and finding resources is easy but using the Raspbian operating system isn't easy. We achieved 80 percent success using dtw. But we don't have a good controller for wrong commands, which can cause some problems. It is not a good way to analyze the signals of the remote and send it again to control the TV. It does not provide a high accuracy rate. We had a success of 60 percent. It might make more sense to control the device using the internet for smart TVs, but the IR system gives us flexibility to control non-smart TVs. We designed the system to work without using the internet. This provides us a great deal of security. Also our system is controlled using turkish commands.

6.2. Hardware

Hardware devices are more expensive in other smart home systems. Usually wifi relay is used. These relays are designed for lamps such as table lamps, not room lamps. You cannot turn your lamp on or off using both the lamp switch and the voice command. Using a relay is a good way to turn on or off the lamp or other electronic devices. It provides less cost. In other smart home systems, you must make either wifi or bluetooth connection to control the devices. By using an IR receiver and transmitter, you can adapt any device that works with the remote controller to your smart home system.

6.3. Cost

The cost of the smart home system we have installed is approximately 350 pounds. The most expensive equipment in our system is Raspberry Pi. Google home mini price is between 300-350 liras and Amazon echo 350-400 liras. When you buy these devices, you only get a device that you can talk to. To control the lamp or other electronic devices in your home, you need to buy extra and expensive equipment. We established our system at a lower cost than these systems and to control some electronic devices. Our system is cheaper than other smart homes, including control equipment such as relays and IR transmitters.

7. CONCLUSION

In this project, we aimed to develop a smart home system on Raspberry Pi. Our goal is to manage our house by using voice commands. For this we have written a code that can understand voice commands. We used DTW and MFCC in our design. We installed a Rasbian operating system on our Raspberry Pi. Then we implement our code on Raspberry Pi. Raspberry Pi has general purpose input output pins to perform commands. We used 2 channel relays to open and close the lamp. When we say turn on or off lamp, the raspberry pin gives us output signal and this open or close lamp using relay. Also you can open and close this lamp by normal switch. To do this we used 2 switch lamp circuits and did research about this circuit. This circuit is called Vavien. Then we did Tv remote controller. To do that we used the lirc package, IR receiver and transmitter. We decoded tv remote controller signals by using a

receiver and lirc package. Lirc makes an IR signal file. Then we used a transmitter to send these signals. According to our command for Tv we send these signals by transmitter. Using the lirc package and decoding these signals are very hard because the Lirc package is an old technology. Tv doesn't receive signals sent from the transmitter very well. This may be a transmitter-induced problem. Current movements in Raspberry may also prevent these signals from being sent correctly. We encountered many problems because it's hard to use Raspberry pi's operating system. Connecting the microphone to Raspberry and using the lirc package are one of them. We added a hotword detection system for this project. We used a snowboy package. We choose our hotword as Jarvis and we send some Jarvis voices to train this system. This system uses voice activity detection and deep learning technologies. Then we added this to our code. Now our system is always listening to us and when we say Jarvis it gets active and operates our commands. This hotword detection system works offline. So we don't get data from our users. This does not create problems for ethical responsibilities.

7.1. Learning Strategies

To do this project, you need to have software and hardware information. We have acquired some software and hardware information in our university life. We used this information in our project and while doing research. We examined the systems that exist in the market. We researched how we can make them cheaper and in the style we want. We adapted to the Python language using our previous software information. In the hardware part, we used the electrical and signal information we learned in our engineering education. While researching, we got a lot of information from the internet. Because these smart home systems are not new systems and they are the systems used. We tested the accuracy of the information we found and blended it with our own information.

7.2. Professional and Ethical Responsibilities of Engineers

Engineers should always be honest with humanity. This is true for every person, but failure of engineers can lead to much more serious and bad consequences. Smart home systems can bring with it many security risks. These systems can be hacked because they are installed depending on the internet. The installed system can listen to you continuously and record its data. This is an unethical behavior. While doing this project, we acted in accordance with ethical values. We designed our system so as not to save personal data. We did not abuse people's personal information. When installing our system, we attached importance to human security. We eliminated the possibility of hacking because it is a local system. We created a project ourselves without stealing what others are doing in terms of software and hardware.

7.3. Further Discussion

We worked on the internet of things while doing our project. Smart houses will take an important place in our lives in the future. Because these systems have important effects on the energy efficiency and safety of the houses. Systems that turn off the lamp when you leave the room and automatically turn off the fountain when you leave it on, save energy, and systems that report the necessary locations in case of fire or theft increase safety. At the same time, these systems are a good target for those who want to get your information. In the future, these systems will evolve from smart homes to smart cities and will require the use of larger systems and systems will be developed with the data obtained from them. We also designed a voice assistant in our project by working on speech recognition. We aimed to manage your home using a voice assistant. Voice assistants and speech recognition technologies are among the most important technologies of today and tomorrow. They have many different uses and make our lives easier. These voice assistants understand what you say and follow your commands. Thanks to these technologies, things will be automated and the need for manpower will decrease.

REFERENCES

[1] Cannistra M.,''Fully Accessible Guide to Smart Home Tech for the Disabled and Elderly'', May 23 201

[2] Al Smadi, Prof-Takialddin, "Automatic detection technique for voice quality interdisciplinary methodologies", Journal of Advanced Sciences and Engineering Technologies(JASET). 1. 1-6. 10.32441/jaset.v1i1.54. , 2018

[3] G. Diğken and T. İbrikçi, "Recognition of non-speech sounds using Mel-frequency cepstrum coefficients and dynamic time warping method," pp. 144-147, doi: 10.1109/SIU.2015.7130277, 2015

[4] Stan Salvador, and Philip Chan. "FastDTW: Toward accurate dynamic time warping in linear time and space." Intelligent Data Analysis 11.5 (2007): 561-580.

[5] Guoguo Chen, 'Snowboy, a Customizable Hotword Detection Engine', 2016 https://snowboy.kitt.ai/docspartials/docs/index.html

[6] https://www.wikizeroo.org/index.php?q=aHR0cHM6Ly9lbi53aWtpcGVkaWEub3JnL3dpa2kvU3BlZWNoX3JlY29nbml0aW9u

[7] Pfundstein, Georg. "Hidden Markov Models with Generalised Emission Distribution for the Analysis of High-Dimensional, Non-Euclidean Data." 2011.

[8] https://www.wikizeroo.org/index.php?q=aHR0cHM6Ly9lbi53aWtpcGVkaWEub3JnL3dpa2kvU21hcnRfaG9tZV90ZWNobm9sb2d5

[9] Pankaj Khatri, 2-Way Light Switch, Retrieved from;

https://circuitdigest.com/electronic-circuits/2-way-light-switch

[10] http://www.piddlerintheroot.com/ir-blaster-lirc/

[11] The Python 2.7 libraries retrieved from; https://docs.python.org/2.7/library/
The Python 3.8 libraries retrieved from; https://docs.python.org/3/library/

[12] Raspberry Pi Snowboy Hotword Detection, Retrieved from;

https://pimylifeup.com/raspberry-pi-snowboy/

[13] Easy Setup IR Remote Control Using LIRC for the Raspberry PI, Retrieved from;

https://www.instructables.com/id/Setup-IR-Remote-Control-Using-LIRC-for-the-Raspbe

[14] The raspberry Pi 3 B retrieved from ;

https://www.raspberrypi.org/products/raspberry-pi-3-model-b/

[15] 2 Channel 5V Relay Module, Retrieved from;

http://wiki.sunfounder.cc/index.php?title=2_Channel_5V_Relay_Module

[16] The IR transmitter retrieved from; https://components101.com/ir-led-pinout-datasheet

[17] The datasheet of IR transmitter retrieved from;

https://pdf.direnc.net/upload/tsop34136-ir-alici-modulu-datasheet.pdf

[18] The datasheet of transistor retrieved from;

http://www.unisonic.com.tw/datasheet/2SC945.pdf