A Novel Approach for Smart Home Using Microsoft Speech Recognition

Palsania Muzaffar
Student, Dept. Of Electronics Engg
Shah &Anchor Kutchhi Engineering College
Mumbai, India

V.C.Kotak
Associate Professor, Dept of Electronics Engineering
Shah &Anchor Kutchhi Engineering College
Mumbai, India

Abstract — **The era of home automation is leaping forward towards the future in creating ideal smart home environments. However, home automation is also developed specially for those who need special attention like elderly patients, sick patients and handicapped person. Thus, different control schemes are adopted to help them do daily routine. In this paper, we present a smart home automation system using voice recognition. The scope of this research work will include the control and monitoring system for home appliances from Graphical User Interface (GUI) using Microsoft Visual Basic software that use Microsoft Speech Recognition engine as an input source and being control wirelessly using ZigBee and at the receiver side, signal is recognized and processed by the microcontroller and the required action is been addressed by Switching the appliances ON or OFF . The research methodology involved is application of knowledge in the field of radio frequency communication, microcontroller and computer programming.**

Index Terms— **Home Automation, Voice Recognition, Microsoft Speech Recognition, Microsoft Visual Basic.**

I. INTRODUCTION

The demography of the world population shows a trend that elderly population worldwide is increasing rapidly as a result of the increase of the average life expectancy of people [1]. Home automation is one of the fast growing industries that keep promising and satisfy the world population in such many ways. It been created due to many aspect such for those who seeking luxury modern lifestyle while others being offers to those with special needs like elderly and disable person [2].
"Home automation is a very promising area. Its main benefits range from increased comfort and greater safety and security, to a more rational use of energy and other resources, allowing for significant savings. It also offers powerful means for helping and supporting the special needs of people with disabilities and, in particular, the elderly. This application domain is very important and will steadily increase in the future [3]."
Due to its potential, several products has been design and commercialized for helping and supporting the special needs of people with disabilities and, in particular, the elderly. One of the products available at the market is known as *uControl* with 7 inch touch screen panel [4]. This system come with the ability to control home appliance via cell phone, Wi-Fi, email, SMS and even GPRS giving the wider selection to the users.
Many of the commercial products use remote control whether it has button or fully touch screen. Still, monitoring and controlling the appliances need some movement and physical contact. Thus, this will be a burden to disable person especially for the disabled and elderly people. As for this project, the proposed solution is to develop a wireless remote control for the home appliances which can be controlled using voice. This system include Graphical User Interface (GUI) features to guide the user which also using voice. The aim of this paper is to develop a device control home appliances via human voice. Microsoft Speech Recognition engine will be used as voice capture device. The Graphical User Interface (GUI) will be created using Visual Basic.

There are several objectives involved in this project that should be focused in order to achieve the design of the project. The idea is to create smart home systems that use biometric method such as human voice as directive to activate electrical appliances. Hence, voice will be used as input to the system. The idea is to design a simple, yet friendly Graphical User Interface (GUI) to aid users especially disabilities and elderly person to do their daily home routine. The system should be using a simple understanding language for easy guidance. The idea is using the wireless system to control home appliances wirelessly, which provide easier installation rather than heavy reconstruction by using wired system. The idea is to analyse the training procedure and overcome the inaccuracy to achieve higher percentage.
This present paper is organized as follows. In Section II we will discuss the human speech characteristics including overview of speech recognition using Microsoft Speech SDK. Section III provides an overview of graphical user interface (GUI) for Home Automation which include software and hardware implementation.

II. OVERVIEW OF SPEECH RECOGNITION

A. *Speech Characteristics*

The speech signal created by the vocal cords, travels through the vocal tract, and produced by the speaker

mouth. Speech signal can be divided into sound segments which have some common acoustic properties for short term interval. It has two major classes which are vowels and consonants. Generally, speech signal can be voiced (periodic source) or unvoiced (aperiodic source). Usually for vowels sound are voiced while for consonants it can be voiced or unvoiced. The sound such as *s, f, sh* is unvoiced and usually consider as a noise in digital signal processing [5].

B. Speech Recognition

Speech recognition is the process of converting an acoustic signal to a set of words. The applications include voice commands and control, data entry, voice user interface, automating the telephone operator's job in telephony, etc. They can also serve as the input to natural language processing.

There is two variant of speech recognition based on the duration of speech signal which is Isolated Word Recognition (IWR) and Continuous Speech Recognition (CSR). IWR defines by each word is surrounded by some sort of pause while CSR with words run into each other and have to be segmented, making it a sentence [6].

Speech recognition is a difficult task because of the many sources of variability associated with the signal such as the acoustic realizations of phonemes, the smallest sound units of which words are composed, are highly dependent on the context. Acoustic variability can result from:

- Changes in the environment
- Changes in the position and characteristics of the transducer
- Changes in the speaker's physical and emotional state, speaking rate, or voice quality

Such variability is modeled in various ways. At the level of signal representation, the representation that emphasizes the speaker independent features is developed.

C. Microsoft Speech SDK 5.1 (SAPI 5.1)

Microsoft Speech SDK is a software development kit for building speech engines and applications for Microsoft Windows.[7] Designed primarily for the desktop speech developer, the SDK contains the Microsoft® Win32®-compatible speech application programming interface (SAPI), the Microsoft continuous speech recognition engine and Microsoft concatenated speech synthesis (or text-to-speech) engine, a collection of speech oriented development tools for compiling source code and executing commands, sample application and tutorials that demonstrate the use of speech with other engine technologies, sample speech recognition and speech synthesis engines for testing with speech enabled application, and documentation on the most important SDK features.

D. SAPI Overview

The SAPI application programming interface (API) dramatically reduces the code overhead required for an application to use speech recognition and text -to-speech, making speech technology more accessible and robust for a wide range of applications.

The SAPI API provides a high level interface between an application and speech engines. SAPI implements all the low level details needed to control and manage the real time operation of various speech engines.

The two basic types of SAPI engines are text-to-speech (TTS) systems and speech recognizers. TTS system synthesizes text strings and files into spoken audio using synthetic voices. Speech recognizers convert human spoken audio into readable text strings and files.
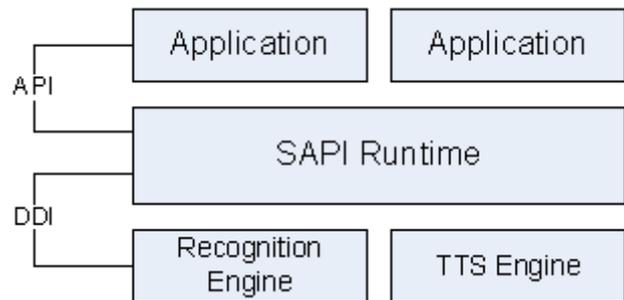

Fig. 1.SAPI Flow Overview

III. DESIGN METHODOLOGIES

In this section, we will describe the design details of our home automation.

A. Software Implementation

Microsoft Visual Basic software is used as monitoring and control section. The system start by initialized the available com ports that has been ready. This available com ports availability is displayed in drop down list in a combo box object. The availability of the com ports depend on the serial ports that been connected to the PC. For this project, a com port is used. As the GUI start, *timer1* function object also starts synchronously. Next, as the com port been selected, serial port will be opened by clicking the connect button to ensure the software and hardware connection. The program will display an error message along with voice guided that tell the user that the com port is not yet open or been selected. If the com port is open then the *timer1* function object will generate 'tick' event for transmitting data to the end device.

For software design, the GUI control function comes with two choices for the user which is by using "voice control" as the primary option or "click button" as secondary option. Both options perform the same function. As the command being given, the voice of SAPI engine will reply as to tell the user that the command has been executed. Hence, user will know that the commands they have given are performed well. Figure 2 shows the flowchart of software algorithm for this project.
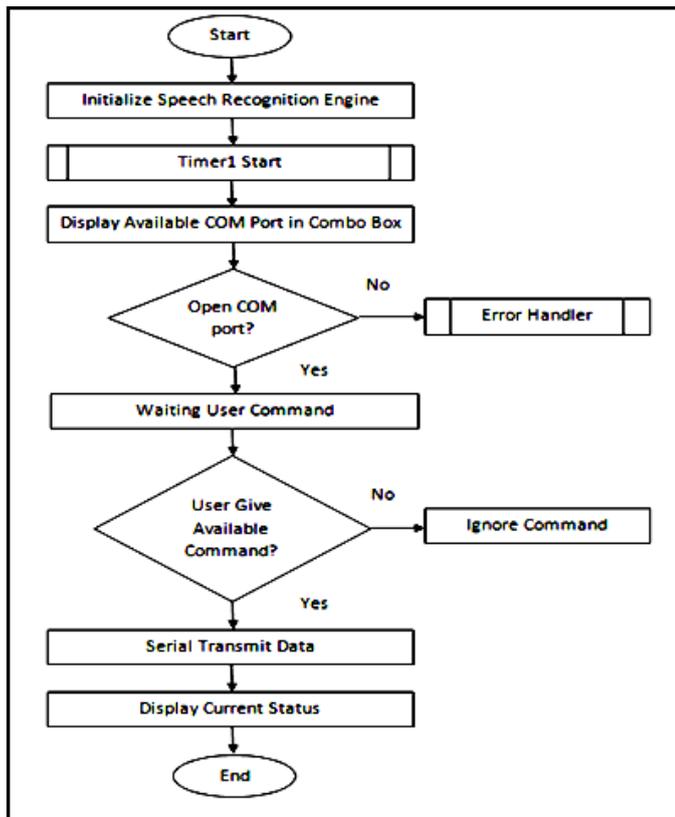
Fig. 2 (a) Main Flowchart of Home Automation System

B. Hardware Implementation

In order to demonstrate our system operation in real-life implementation, we create a simple prototype that represents home appliances. Microcontroller is used for controlling the end device at the receiver that control home appliances.

Command from user that receives from personal computer (PC) will send the data to the microcontroller using ZigBee protocol. Figure 3 below shows the overview of the system.
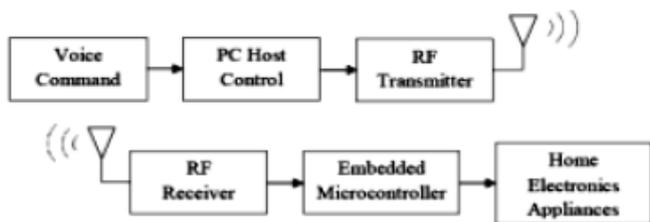

Fig. 3. System Design Overview

In this project, there are 8 outputs which include DC mode. For DC model, LED will be used to represent the appliances in the house. When data is sent from host PC, the codes used to represent the command were in ASCII. ASCII was established to achieve compatibility between various types of data processing equipment making it possible for the

Components to communicate with each other successfully. This code is representing in byte which in binary, decimal, or hexadecimal.

For example, when "A" is sent, the code is converted into byte, for this case we assume in decimal, which are 65. When PIC received 65, it will process the code and convert it back to "A". Hence, the next process will be executed based from the code received.

Command	Mode	Serial Communication Code
CHECK FOR TIME FOR NOW	Notify user the exact time at that moment	N/A
PLAY MUSIC	Play the music from music list	N/A
VOLUME INCREASE	Increase the music volume	N/A
CHECK STATUS FOR KITCHEN HALL	Notify user the kitchen hall current status	N/A
SWITCH ON LIVING HALL LIGHT	Turn ON living hall light	A
SWITCH OFF LIVING HALL FAN	Turn OFF living hall light	D
FAN FULL SPEED	Turn the fan to maximum speed	E
SWITCH OFF KITCHEN HALL LIGHT	Turn OFF kitchen hall light	K
ACTIVATE LIGHT ON	Turn ON lamp 1	L
LIGHT ONE DEACTIVE	Turn OFF lamp 1	M

Table 1 : Example of Command List Available

At the receiver part, PIC 16F877A, a 40-pins microcontroller is used to operate the end device system since it has many ports and most importantly is supports UART features [8, 9]. Besides, this PIC acts as the heart for end device unit which execute the command from Visual Basic through serial port communication. The CPU is clocked at 20MHz for full speed operation. A XBee Pro 100mW wireless module also used as the interface between host PC and the end device unit.

IV. ANALYSIS

User need to install Microsoft Speech SDK in order to activate the "speech" mode.

A. System Verification

To start the speech recognition, user need to say "Start Listening" or pressing the blue microphone icon to activate the

speech recognition system. Next, user need to speak certain command to perform specific task. For example, when users say "SWITCH ON LIVING HALL LIGHT" hence the command will be executed and it will transmit the code to hardware system for operation. In Table 1 we listed several command available for our home automation system. As for software, once user give the command the system will replies to tell the user that the command given has been executed such as "LIVING HALL LIGHT IS ON". The code was transmitted in ASCII mode.

B. Improving the Accuracy
Voice pattern are not always stable due to change in mood and environment expression. Thus, sometimes user has to say the same command a few times before it can be recognize. SAPI provide the platform that let user to modify the dictation according to our own ways. Figure 6 shows the speech option that allow user to modify to get better accuracy. There are three options that can be used by user to setting their own option such as adding the new word, prevent word, and even change existing word.

In order to recognize the word pronounce by the user, normally an error happened due to comparison of the suitable word. This problem usually happens when user pronounce the word not fluently or in understandable ways. Besides, each race has its own dialect. For example, Britain and American having two differences style of speaking although both countries speak the same language. This not includes other region such as Asia with majority that not being able to pronounce English correctly. Besides, the similar sound in pronunciation also will lead to misunderstanding. Theoretically, "ONE" should be pronouncing as "WUHN" and "WANT" as "WAWNT". Still, the pronunciations are heard similar to most people. Hence, to improve the dictation accuracy, option in the SAPI need to be used. The "Adding the New Word" option gives the biggest advantages to user. In this option, user can add new word although in different language. Theoretically, speaks creates a waveform. Hence, the technique used by most voice recognition developer is capturing the waveform of the word pronounce. Thus, different language can be recorded and being used.

Another trick to improve the accuracy is learning on how the word being saved. In example, "Hello" and "HELLO" are same words but the phrase recorded follows the word that user decide. Thus, if the word not patterned, two probabilities will rise and can cause an error. This is quite similar to ASCII concepts where "A" and "a" alphabet are same letter but different in numerical numbers. "A" is set to decimal 65 while "a" is decimal 97. Hence, the same concept applied for this project to improve the dictation accuracy. Lastly, "Changing the Existing Word" option can be used to further improve the accuracy level. As discuss in previous paragraph, "ONE" and "WANT" usually being pronounce in same way although both

have different meaning. For this project, word that not being used need to disabled to decrease the percentage of probability, thus increase the percentage of accuracy.

V. CONCLUSIONS AND FUTURE WORKS

This paper present the design of a home automation system with computer monitoring based on speech recognition and networking control feature. Speech recognition able to ease people especially those who having disabilities such as vision impaired and also elderly. Although nowadays many developer creating smart home system, only a few using voice to access the control centre. With SAPI 5.3, the result of the recognition is reliable with lowest error occurred. The use of wireless communication makes the system more convenient and easy to install instead using wired communication. Besides, it provided low cost power consumption.

As for recommendation, there are few suggestions that can be considered for further research to improve this project:
1. The monitoring part not only limited to the ON/ OFF the home appliances only.
2. Motion sensor may add for automatic lighting and turning ON the fans in the area where user were there.
3. Schedule may add to enable user to set the ON/ OFF timer for home appliances.

REFERENCES

[1] N Keyfitz, W Flieger, "World population growth and aging: Demographic trends in the late twentieth century" University of Chicago Press, 1990.

[2] Khusvinder Gill, Shuang-Hua Yang, Fang Yao, Xin Lu "A Zigbee-Based Home Automation System" 2009. 55(3), pp. 422-430.

[3] Renato J. C. Nunes, Jose C. M. Delgado, "An Internet Application for Home Automation", 10th Mediterranian Electrotechnical Conference. MEleCon 2000, vol 1, USA: IEEE 2000, pp. 298-301.

[4] (2010) Home Automated Living website. [Cited 2010 14th Oct]. Available: http://www.homeautomatedliving.com/default.htm.

[5] Ze-Nian Li, Mark S. Drew, "Fundamental of Multimedia", Pearson Prentice Hall, 2004, pp. 379.

[6] Adam L. Buchsbaum, Raffaele Giancarlo, "Algorithmic aspects in speech recognitions: an introduction", Journal of Experimental Algorithmics (JEA) Volume 2, 1997, Article No. 1.

[7] Hao Shi, Alexander Maier, "Speech-enabled windows application using Microsoft SAPI", IJCSNS International Journal of Computer Science and Network Security, VOL.6 No.9A, September 2006, pp. 33-37.

[8] John B. Peatman, "Embedded Design with the PIC18F452 Microcontroller", Pearson Education, 2003.

[9] Microchip Technology Inc. *PIC16F877A Data Sheet*, United States of America: Microchip Technology Inc. 2009.