Smart Home with Biometric System Recognition

Nor Syazwani Md Noh¹, Haryati Jaafar¹, Wan Azani Mustafa¹, Syed Zulkarnain Syed Idrus², A. H. Mazelan³

¹Department of Electrical Engineering Technology, Faculty of Engineering Technology, Universiti Malaysia Perlis, UniCITI Alam Campus, Sg. Chuchuh, 02100 Padang Besar, Perlis, Malaysia.
²Center of Excellence Geopolymer and Green Technology, Universiti Malaysia Perlis, 01000 Kangar, Perlis, Malaysia.
³Sport Engineering Research Centre (SERC), School of Manufacturing Engineering, Universiti Malaysia Perlis (UniMAP), Pauh Putra Campus, 02600 Arau, Perlis, Malaysia

syazwani@yahoo.com

Abstract. In the last decade, smart homes have become applicable and widely developed where all the electronic devices and household appliances are controlled from a centralized control unit. However, the rise of misuse such as the security aspect and interoperability mechanism is proving them to be less reliable. Moreover, the traditional home automation system is complex, expensive and they are inconvenient for the aging population or disable persons. Besides, most smart home technologies forcing users to use their hands in order to control electronic devices. Therefore, the smart home with biometric system recognition is proposed to increase security, efficiency and decrease manual labor effort. This project used biometric of the human voice as an input command. The scope of this project is only focused on fan and lamp as an output of home appliances. The proposed system is divided into two main parts which are speech recognition (software development) and electrical home appliance control system (hardware development). In the speech recognition, the process contains two main modules which are feature extraction and pattern matching. Meanwhile, in hardware development, the microcontroller board is used to process the home appliance system such that the system can be operated automatically. It is expected that this project allows the user to conveniently control the home appliances by using voice. The experiment shows that the highest performance of the system achieves an accuracy of 83.25% with the number of k, training, and testing is 1, 20, and 40 respectively.

Keywords; smart home, biometric system, signal processing, software development, hardware development

1. Introduction

A smart home is a residence that equipped with technologies including internal network, intelligent control, and home automation. It is anticipated and responded to the needs of the users or occupants. In addition, it becomes applicable and widely developed where all the electronic devices and household appliances are controlled from a centralized control unit [1-3]. Nevertheless, the rise of misuse such as the security aspect and interoperability mechanism is proving them to be less reliable. [4-6].
In order to overcome the mentioned problems, the smart home with biometric system recognition has been introduced and widely employed over the years. This new system can work automatically with the aid from biometric recognition for authentication [7,8]. It can identify or verify a user based on a specific behavioral or physiological feature a user possessed. It is considered the best alternative to the traditional smart home system as it increases security, efficiency, and decrease manual labor efforts.

The application of biometric recognition system based on the image (fingerprint, finger vein and face) and signal (voice and speech) have been widely used in the smart home system. Both of the systems are well-performed in term of accuracy and time processing. However, it is found that the biometric system based on an image is more expensive compares to the signal. This is due to the cost of equipment and complexity development of the system [7,8]. In contrast, the biometric system based on signal processing is more convenient, less complex and low cost. Moreover, the voice is easy to be produced whether the user is a healthy or sore throat. Motivated with the effectiveness of this system, this study proposed the smart home system that used biometric of the human voice as an input command to control the appliances.

The system is divided into two main parts which are speech recognition (software development) and electrical home appliance control system (hardware development). For speech recognition, the data is based on the online and offline system. The speech data is firstly acquired from subjects of male and female between the ages of 22 and 24 years old. The speech data is recorded in a clean environment with a different type of voice condition such as normal, fever and sore throat. The collected data is then stored in a database before the identification system is implemented to evaluate the speech signal of the online user. The Matlab program is applied to develop the biometric system for smart home appliances.

Meanwhile, for hardware development, the Arduino UNO is used as a microcontroller board to process the home appliance system. The development of a smart home is a focus for home appliances such as fan and lamp located in the living room, kitchen and bedroom spaces. The overall architecture of this project is illustrated in Figure 1. The system is divided into three phases which are the input voice produced by the user, speech recognition process by Matlab programming, and hardware interface.

![Figure 1. The architecture of a smart home with biometric system recognition](image)

2. Methodology
In the smart home with the biometric system, the process is divided into two main parts which are software and hardware development process. The software development involves data acquisition, audio signal processing, and Graphic User Interface (GUI) development. Meanwhile, hardware development involves prototype design and circuit construction.
2.1 Data Acquisition

The human voices are collected from 20 subjects which are ten males and ten females between the ages of 22 and 24 years old. The total of 1200 speech data is recorded, with 60 speech data collected from each subject. The speech data is acquired in different types of voice condition such as normal, sore throat, and fever. For each type of voice condition, 10 speech data of ‘on’ words and 10 speech data of ‘off’ words are recorded. This means each subject has a total of 20 speech data for each condition. The recordings are carried out by using Sony Stereo IC Recorder ICD-AX412F supported by a Sony electric condenser microphone 32-bit, and 32kHz sampling frequency with the mp3 format. The length for each recorded speech data is approximated 2s. Fig. 2 shows an example of recorded data.

![Figure 2](image)

**Figure 2.** The plot of the speech signal in the digital waveform

2.2 Audio Signal Processing

In audio signal processing, each of the collected speech data undergoes a series of signal processing which includes pre-processing, signal segmentation, feature extraction and classification. The pre-processing involves the process of digitization, filtering, windowing and framing [9]. Subsequently, the signal segmentation is employed to separate the undesired and desired signals. Here, time domain techniques which are energy (E) and zero crossing rate (ZCR) are used [10, 11]. In E technique, it assumes that speech is louder than the background noise. Hence, the energy is high-energy frames to speech and low-energy frames to noise. Meanwhile, the ZCR is the rate where the signal changes from the positive to negative values. Therefore, when the voice is detected, the value of ZCR is relatively low as the higher one is detected as an unvoiced. Fig. 3 shows an example of E and ZCR technique. The red line is indicated the starting point and ending point of E and ZCR. It can be seen, the combination of these techniques able to remove the undesired signal such that the sequencing process is only focused on the main part of the signal.
Figure 3. Speech segmentation using the E and ZCR techniques

After the speech signal is segmented, the data is extracted to obtain the important data and this process is called feature extraction. In this process, the well-known technique name mel-frequency cepstral coefficients (MFCC) is applied. MFCC has selected due to the features are robust to noise which is suitable to be implemented in an outdoor environment that contains interference of background noises [12]. Moreover, calculation of MFCC is based on the human auditory system aiming for artificial implementation of the ear physiology assuming that the human ear can be good speaker recognizer too. The operation of MFCC is shown in Fig. 4. There are 12 mel cepstrum coefficients, one log energy coefficient and three delta coefficients per frame have been set in the experiments [14].

![Typical MFCC process](image)

Figure 4. Typical MFCC process

In order to determine the performance of results, k-nearest neighbor (kNN) is used in the pattern machine process [13]. In this process, the query point is assigned to the class by calculating the distances of the training samples from the query point. The distance between the query point and the training samples is written as in (1).

\[ d(y, x_j) = \sqrt{(y - x_j)^T(y - x_j)} \]  

where \(d(y, x_j)\) is the Euclidean distance, \(x_j\) is the training sample and \(y\) is the query point. The distances are consequently arranged in ascending order and the kNN is selected based on the training sample that has the smallest distance to the query point.
2.3 Graphic User Interface Development

The Graphic User Interface (GUI) is applied to display the online speech recognition systems that identify/verify the real-time user and platform to control the home appliances. It acts as a control or monitoring system that displays the user speech signal, result of identification/verification, and status of homes appliance (on/off) as shown in Fig. 5.

There are two main push buttons namely record and identify the button. Each of this button has its own functions and executes specific processes when activated by the user. When the record push button is pressed, the microphone recorded the user’s speech for 5 seconds and then it stops. After the speech is recorded, the signal is extracted and displayed in the GUI box. Then, the identify push button is pressed to display the result of the user’s identification or the verification as well as the status of home appliances (on/off).

![Figure 5. Interaction platform for an online speech recognition system](image)

2.4 Hardware Development

The smart home hardware system will be working automatically and will be responding to the change of the input condition. In order to develop the system, there are few electrical components have been used i.e. Arduino UNO as microcontroller, infrared (IR) sensor, direct current (DC) motor, light emitting diode (LED), and Xbee in order to develop the system. Fig. 6 shows the full circuit of hardware components and Fig. 7 shows the workflow of the circuit operation.

![Figure 6. The connection circuit of hardware components](image)
3. Experimental Results
The experiments are implemented using Matlab R2016(b) and have been testing in Intel Core i5, 2.1GHz CPU, 2G RAM and Window 8 operating system. The feature dimension size was fixed at 4096 (64×64). For the classification, the number of k is set to 5.

3.1 Performance Evaluation for Offline System
In order to evaluate the performance based on signal segmentation, the analysis test of the signal without segmentation and signal with segmentation is carried out. The number of training and testing is fixed to 20 and 40 respectively. The performance is evaluated based on classification accuracy (CA). The evaluation measures are based on the CA such that:

![Figure 7. Circuit operation workflow](image-url)
\[
CA = \frac{N_c}{N_A} \times 100\%
\]  
(2)

where \(N_c\) is the correct identified number of samples and \(N_A\) is the total number of test samples.

### Table 1. Segmentation Result

| Condition      | Manual segmentation | E and ZCR techniques |
|----------------|---------------------|----------------------|
|                | ON CA (%) | OFF CA (%) | ON CA (%) | OFF CA (%) |
| Normal         | 174   87 | 165  82.5 | 186   93 | 182   91 |
| Sore throat    | 168   84 | 138   69 | 174   87 | 168   84 |
| Fever          | 127   63.5 | 119  59.5 | 163   81.5 | 165   82.5 |
| Total          | 469   78.17 | 422  70.33 | 523   87.17 | 515   85.83 |

The results in Table 1 show the combination of E and ZCR outperforms than manual segmentation with total CA of 87.17\% and 85.83\% for ON and OFF words. The results also indicate the CA is outperformed during normal condition. This is because, during the segmentation process for both sore throat and fever, the E and ZCR techniques are mistakenly segmented the noise part and the background noise is slightly higher than the desired speech.

Consequently, the performance of the speech recognition system based on a different number of testing was investigated. In this experiment, the number of testing were set with 20, 30, and 40. The number of nearest neighbors, \(k\) and training are fixed to 1 and 20 respectively. The performances accuracy of a different number of testing is shown in Table 2. From the result, it shows that 40 number of testing outperforms with CA 85.63\%, better than others.

### Table 2. Performance CA based on a different number of testing

| User | Number of testing |
|------|-------------------|
|      | 20    | 30    | 40    |
| 1    | 60    | 83.33 | 82.5  |
| 2    | 80    | 70.00 | 87.5  |
| 3    | 70    | 66.67 | 85    |
| 4    | 80    | 90.00 | 82.5  |
| 5    | 65    | 73.33 | 85    |
| 6    | 75    | 83.33 | 87.5  |
| 7    | 60    | 63.33 | 90    |
| 8    | 85    | 66.67 | 80    |
| 9    | 95    | 66.67 | 72.5  |
| 10   | 90    | 70.00 | 70    |
| 11   | 55    | 73.33 | 80    |
| 12   | 50    | 76.67 | 85    |
| 13   | 70    | 86.67 | 80    |
3.2 Experimental Results for Online

In the online system, 20 users have been tested where the speech from four users have been processed earlier and saved as the training data. Meanwhile, the other user is considered as untrained user where there was no record of this user earlier. To start the identification process, the user was asked to record the voice instruction (input) by pressing the record button. The recorded signal voice is displayed in the GUI as in Fig. 8. The recorded signal voice is considered as testing data and will be segmented and extracted before the identification process is employed. If the user is successfully identified, the identity of the user will be displayed as shown in Fig. 9. The green colour of the output fan and lamp in GUI indicates that the smart house is activated. However, if the system does not recognize the voice signal of the user, it will prompt an error window appears as shown in Fig. 10. The error dialog indicated the intruder is detected. In this case, the smart house will be deactivated.

To reduce energy consumption, the system will be OFF if the IR sensor does not detect the presence of the user. Then, the message box will be displayed to notify that no user is detected as shown in Fig. 11. Here, the red colour of output fan and lamp in GUI indicate that it is in turn OFF state.

![Image](image-url)

**Figure 8.** The input before the identification process
Figure 9. The result after the identification process is succeeded

Figure 10. Example of error detection for untrained used

Figure 11. Example of displayed result when user absent
4. Conclusion

In this study, the smart home with the biometric system has successfully developed a smart home with biometric system recognition. The system is divided into software and hardware development. In software development, the speech data undergoes a series of signal processing which includes pre-processing, signal segmentation, feature extraction, and classification. By acquiring speech data in different conditions such as normal, sore throat and fever, the identification process has been obtained with the accuracy is more than 85% for the offline system. On the other hand, the GUI is developed and integrated with hardware prototype. The results show that the online system is able to identify the registered user or intruder. The hardware is successfully operated based on the result of the offline system.

Reference

[1] T. Mantoro, M. A. M. Adnan, and M. A. Ayu, “Secured Communication between Mobile Devices and Smart Home Appliances,” 2013 Int. Conf. Adv. Comput. Sci. Appl. Technol., pp. 429–434, 2013.
[2] V. S. Gunge, “Smart Home Automation: A Literature Review,” no. Rtdm, pp. 6–10, 2016.
[3] R. Jayalakshmamma, P. V Naganjaneyulu, and K. Babulu, “Implementation of Integrity of Voice and Face Recognition for Home Security By Using Gsm and Zigbee,” no. 4, pp. 1043–1047, 2012.
[4] A. Cyril Jose and R. Malekan, “Smart Home Automation Security: A Literature Review,” Smart Comput. Rev., vol. 5, no. 4, pp. 269–285, 2015.
[5] F. R. Ishengoma, “Authentication System for Smart Homes Based on ARM7TDMI-S and IRIS-Fingerprint Recognition Technologies,” Int. J. Inf. Eng. Electron. Bus., vol. 6, no. 6, pp. 64–69, 2014.
[6] N. A. Wahile, P. D. Hatwar, and I. M. Padiya, “Home Automation System Based on Speech Recognition,” vol. 5, no. 3, pp. 130–136, 2017.
[7] A. M. Soe, M. M. Latt, H. M. Tun, and Z. M. Naing, “Electronic Control System Of Home Appliances Using Speech Command Words,” vol. 4, no. 6, pp. 323–329, 2015.
[8] S. Asha and C. Chellapan, “Biometrics: An Overview of the Technology, Issues and Applications,” Int. J. Comput. Appl., vol. 39, no. 10, pp. 35–52, 2012.
[9] A. K. Sharma, A. Raghuvanshi, and V. K. Sharma, “Biometric System- A Review,” vol. 6, no. 5, pp. 4616–4619, 2015.
[10] D.A. Ramli, and H. Jaafar, “Peak Finding Algorithm to Improve Syllable Segmentation for Noisy Bioacoustic Sound Signal,” Procedia Computer Science, vol. 96, 2016, pp.100-109.
[11] Y. Tian, Z. Wang, and D. Lu, “Nonspeech segment rejection based on prosodic information for robust speech recognition,” IEEE Signal Processing Letters, 2002, vol. 9, no. 11, pp. 364–367.
[12] C. Guo, R. Zhi, M. Fan, and K. Liu, “Research on voice activity detection in burst and partial duration noisy environment,” International Conference of Audio, Language and Image Processing (ICALIP), 2012, pp. 991-995.
[13] H. Jaafar, N. Mukahar, and D.A. Ramli, “A methodology of nearest neighbor: Design and comparison of biometric image database,” IEEE Student Conference on Research and Development (SCoReD), 2016, pp 1–6.