Building Smart House based on Speech Detection and Recognition System

Kawther Thabt Saleh  
Department of Computer Science, College of Education, Mustansiriyah University, Baghdad, Iraq.  
E-mail: kawtherthabt@uomustansiriyah.edu.iq

Safa Ameen Ahmed  
Department of Chemical Engineering, University of Technology, Baghdad, Iraq.  
E-mail: Safa.a.ahmed@uotechnology.edu.iq

Haitham Salman Chyad  
Department of Computer Science, College of Education, Mustansiriyah University, Baghdad, Iraq.  
E-mail: dr.haitham@uomustansiriyah.edu.iq

Received October 02, 2021; Accepted December 22, 2021  
ISSN: 1735-188X  
DOI: 10.14704/WEB/V19I1/WEB19342

Abstract

Technology in all areas has entered into life in homes (achievement of housework), streets and work, in order to facilitate life matters and saving effort and time, also, the technology has entered the areas of entertainment and communication and achievement of works even remotely and achievement of work that needs strength and short time. Also, artificial intelligence has entered various applications to improve its performance; this research will offer one of the smart technology applications, which is the smart home where the instructions can be handled and performed through speech, in order to add more facilities to our lives and help some people (e.g., elderly or disabled people who live alone) who have difficulty to move and handle the regular appliances in the house. And due to the lack of voice command recognition research in Arabic has resulted in making It difficult to stratify smart house voice command services, especially in the Middle East. Speech recognition is also a nontrivial task in the processing of natural language; this proposed smart house system is based upon detection and recognition of the speech in the Arabic language. Arduino programming was used to accomplish this system.

Keywords

Artificial Intelligence (AI), Smart House, Speech Recognition, Arduino, Arabic Language.
The "Smart House" term is commonly utilized for referring to any setting that has been designed to assist people in their daily lives to encourage an independent life (Winkler 2002; Helal et al., 2003; Chan et al., 2008; Sanchez et al., 2017). Smart technology is used in several specialties and is carried out for helping some applications, such as medicine, building, security, manufacturing, etc. (Dlodlo et al., 2012; Ahmed et al., 2018). Also, this technology is used to build many smart home systems. It comes in several forms, for example, Park et al., 2007) presented a novel robotic smart house (Intelligent Sweet Home) that has been improved in Korea at KAIST; it is based upon many robotic agents and has the aim to test innovative concepts for the elderly and people who have disabilities to help them live independently.

The work has been focused upon technological solutions for human-friendly mobility\motion assistance and advanced human-machine interfaces which make both robotic assistive systems and home-installed utilizations easy to operate. The design of the smart house involves an intelligent bed, a smart wheel-chair, and a robotic hoist to quickly move the user from the wheelchair to the bed and vice versa. The architectural elements conform to the majority of users' requirements and recommendations; these guidelines were obtained from a special questionnaire for persons with disabilities. The smart house responds to the user's instructions as well as to the needs recognized by the user. Different interfaces have been researched and tested depend on hand movements, speech, form language, and position (Park et al., 2007).

Sehili et al., 2012 introduces speech and sound analysis framework assessed in a real smart house environment due to an acquired corpus of data. Signal identification, sound\speech recognition, and classification are the fourth stage of examination. For each stage, the outcomes are described. The first experimentations displayed promising results, whether for the modules individually tested or for the entire system (Sehili et al., 2012).

Desot et al. 2019b suggest an End-to-End (E2E) model to carry out classification directly from the raw speech data. Thus, the E2E method is tailored for this particular role and prevents the propagation of errors. In addition, the E2E model for purpose classification will manipulate prosodic features of the speech signal (e.g., imperative vs. question voice). In a real smart house, tests on a corpus of voice commands that have been obtained show that state-of-art pipeline base-line remains better compared to the E2E method. None-the-less, utilizing artificial data generation methods demonstrated that it is possible to make substantial improvements to the E2E model to achieve competitive
efficiency. This has opened the way for more E2E Spoken Language Comprehension studies (Desot et al., 2019b).

Ranger, Bernard et al. presented an interaction system for a smart home. It has been constructed on a multithreaded, multimodal interchange engine for conversations. The framework offers a user interface in natural language to manage household equipment, appliances, or household features. By detecting devices like a mobile phone, a computer laptop, or a tablet device, the smart home automation agent can collect feedback from users. From inside the household or from distant locations, users communicate with the device. The smart home scheme will obtain data from sensors or any other device with which it is interfaced. The device uses rules for the communication guide to process responses to the input of both user and sensor and drive the intended to answer resulting from such input (Ehsani et al., 2019).

Abu et al., 2020 suggested a system that is referred to as recognize speech and extract feature that has two main processes. Data is obtained by the audio fingerprint in the feature extraction method (the spectral peak command). Audio fingerprint matching can be employed for voice recognition to process audio commands. The result of this scheme is when the user activates or deactivates the target home appliance to provide the voice order. The second result is that the background noise impact of the system is minimized by utilizing the robust quad algorithm, and also, the scheme's precision has been improved. The outcomes of the project have shown that implementing the proposed algorithm on a Raspberry Pi and achieving a high recognition rate of 87 percent is rational. In the presence of 15 dB noise, this suggested algorithm achieved 82 percent of the higher accuracy rate (Abu et al., 2020).

Shostak et al., 2020 presented a method of using a "Smart house" multi agent mechanism scheme of smart components. The framework was based on the example of 3 intelligent robots. The subject independent multi-agent kernel that contains the next elementary components is the main part of the system architecture under improvement: direct access service, agent class library, messaging service, ontology, agent community. It has been discovered that in the context of this issue, the multi-agent method employing ontologies meaningfully exceeds conventional approaches in performance (Shostak et al., 2020).

**ARDUINO**

The IDE source code is issuance within the general public authorization (GNU); the languages C and C++ have been used to support the version of Arduino IDE utilizing
distinct code constructing rules. The Arduino IDE offers a software library that delivers several public output and input performances from the Wiring project. If a programmer wants to write program code, he just needs the main tasks, starting the plan and the main program loop, that are combined and connected to the stub main() program with the GNU tools chain into an implementable cyclic procedural program. Moreover, it contained the IDE distribution. The program avrdude id has been employed by The Arduino IDE in order to change the implemented code into a hexadecimal text file encoding. In the board's firmware, the loader program loads this text file is loaded to the board of Arduino. For more details, you can read the sources (Martín-Ramos et al., 2016; Gingl et al., 2019).

Proposed System

The suggested smart house system has been based upon a detection and recognition speech system—furthermore, Sensitive fire and gas alarm devices for detecting thieves, which is done automatically. Fig. 1 presents the circuit structure of the Smart Home. Fig. 2 shows the Technical Specifications of ARDUINO and the ARDUINO elements.

![Figure 1 Structure of the Smart Home and the connection between its elements](image)

![Figure 2 ARDUINO Technical Specification and the element of ARDUINO](image)
Hardware Components

The proposed system consists of a set of basic hardware components (e.g., Arduino Uno R3, LCD 16x2 display, Voice, Recognition module, MQ-2 Gas sensor, PIR motion sensor, LDR sensor, Servo motor, Fan, Relay, LED, Buzzer). Which were used in this scheme and these components are:

1. ARDUINO UNO R3

Arduino Uno can be defined as a micro-controller board that has been based upon ATmega 328P. Atmel, in the mega AVR family, creates the ATmega328, which is a single-chip microcontroller. In 2016, the Microchip Technology gained Atmel). It has an improved 8-bit RISC Harvard architecture processor core, such as shown in Figure 3 (Faludi 2010). It has fourteen digital pins output/input of which six analogy inputs, six may be employed like PWM outputs, a power jack, a USB connection, an ICSP header, a 16 MHz quartz crystal, and a reconstruct button. Fig. 4 presents all these details [20]. It has everything that is required for supporting the micro-controller. Figure 5 shows the ARDUINO Uno R 3.
2. LCD 16x2 DISPLAY

The Liquid Crystal Display (LCD) is an electronic screen unit that is used in various applications. A 16 x 2 LCD screen is quite a fundamental module. Several devices and circuits use the 16x2 LCD display. These units are favored above seven pieces and extra multi-piece LEDs for the reason that the LCDs are inexpensive; simply programmable; have no limitation of displaying animations, custom and even special characters (different in seven pieces), and so on. The name of 16 x 2 LCD means that it can show 16 symbols for every line, and there are two as lines. Every symbol is shown in a matrix of 5 x 7 pixels using this LCD, which has 2 registers, Data and Command. The instructions of command are specified to the LCD that save in the command register. This command is an order specified to LCD for the purpose of accomplishing a pre-described mission as initializing it, Set the location of the cursor, clear the screen, monitor the display, etc. In the data register, the data that will be shown on the LCD is stored. The data is represented the ASCII value of the symbol to be presented on the LCD (Niranjan et al., 2017). Figure 6 Shows an LCD 16*2 display.

3. Voice Recognition Module

ELECHOUSE Voice Recognition Unit is a board that is built-in and simple regulator speech recognition. The module of the effect is a speaker focused on the identification of voices. It can assist about eighty voice commands in general. It is possible to work at most 7 voice commands simultaneously. Any sound may have been trained as an order. Firstly the users should practice the unit previously let it distinguish any voice order command.
There are two means of managing this board: General Input Pins (function part) and Serial Port (full function). General Output Pins on the board produced many kinds of waves while the corresponding voice command was recognized (Bashra et al., 2018; Jinan 2019). Figure 7 and Figure 8 Show the Voice recognition module.

![Figure 7 Voice Recognition module](image1)

![Figure 8 Voice Recognition module with the microphone](image2)

4. MQ-2 Gas Sensor

SnO₂ is the sensitive physical part of MQ-2 gas device sensor, which cleans the air through the decrease in its accessibility. In the case where the objective adjust gas is present, when the gas concentration rises, the sensor's conductivity also increases (Chalasani and Mohan 2017). Figure 9 displays MQ2 Gas Sensor.

![Figure 9 MQ2 Gas Sensor](image3)
5. PIR Motion Sensor

A photo-resistor also is called a light-dependent resistor, photo-conductive cell, LDR. It is a variable light-controlled resistor. While the incident light intensity increase, the resistance of a photo-resistor is decreased; but, it displays photo-conductivity. A photo-resistor can be used in detector circuits of light-activated and light-sensitive of dark-activated switching (Narayana et al., 2015). Figure 10 presents the PIR motion sensor.

![Figure 10 PIR motion sensor](image)

6. LDR Sensor

Another component is LDR, which has been used in our proposed system. The LDR has a variable resistance that is affected by the light concentration that drops on the LDR. This feature permits them to be employed in light recognizing circuits. The Working Principle of LDR permits these units to depend on the light; the resistance decreases when the light drops upon it and rises in the dark. When the LDR is found in a bright place, its resistance will decrease. On the other hand, when the LDR is found in a dim place, its resistance is increased (Antony et al., 2018). Figure 11 shows LDR Sensor.

![Figure 11 LDR Sensor](image)

7. Servo Motor

It has many characteristics. It's tiny and lightweight, and its output power is high. Moreover, it can rotate about 180 degrees, in other words, 90 in every direction, and its...
mechanisms are similar to the typical types, but it is smaller. In order to control these servos, any servo code, hardware or library can be used. It is better, particularly when it fits in small places and for beginners who want to make materials travel without construction a motor controller with gearbox and feedback (Brian and Derek 2017). Figure 12 presents a Servo motor that was used to open the door and close it.

Figure 12 Servo motor

8. Fan

In this proposed system, the fan unit has been used. A DC motor is employed in the fan and any type of rotational electrical device that transforms direct existing electrical energy into mechanical energy. The general popular kinds trust on the requirements that are made by magnetic fields. Practically all DC motor categories have several internal techniques (i.e., electronic or electro-mechanical), which are used for the sporadic alteration of the direction of present movement in the motor part (Cornel and George 2018). All of that has been illustrated in Figure 13, which presents the fan.

Figure 13 The fan

9. Relay

A relay may be defined as an electrically operated switch. An electromagnet is used by many relays, where u is a mechanically run switch. On the other hand, the functional
principles are utilized as well, for example, solid-state relays. Fig. 14 presents Relays that are employed where it is compulsory to manage a circuit via a distinct signal of low-power, or where some circuits has to be managed by one signal. Length area telegraph circuits used the first relays as amplifiers: it repeated the arrival signal in/from a circuit and retransmitted it on a different circuit. Relays have been extensively utilized in the early computers and telephone exchanges for the purpose of performing the logical operations. Relays are alterations between two cases (close and open circuits electronically or electromechanically). Relays were opening and closing contacts in another circuit in order to control one electrical circuit. An open communication occurs when a contact relay is usually open (ON). On the other hand, when the relay isn’t energized, then, the contacted relay is not energized (Jabbar et al., 2019).

![Figure 14 The Relay](image1)

10. LED

Light Emitting Diode (LED) is a source of semiconductor light. When the flow streams through the LED, it emits light. In the semiconductor, the electrons remerge with electron pits, so it is casting energy in the photon form (Brian and Derek 2017). Figure 15 shows the led.

![Figure 15 The LED](image2)
11. Buzzer

A beeper, which is also called a buzzer is a machine of audio signalling, which may be electro-mechanical, mechanical, or piezoelectric (piezo for short). Standard beepers and buzzers are used in alarm machines, timers, and verification of user input for instance a keystroke or mouse click (Brian and Derek 2017). Figure 16 shows the Buzzer.

![Figure 16 The Buzzer](image)

Results and Discussion

Smart House is one of the artificial intelligence applications that provide for our life many facilities to save time and effort. Most recent research studies on the recognition of voice commands in English for the smart house (Abu et al., 2020). The absence of voice command recognition systems in Arabic makes it difficult to introduce voice command services for smart homes, especially in the Middle East. Speech recognition is also a nontrivial task in the processing of natural language. In this system, we used speech recognition to use and handle house components. Figure (17:a, 17:b, 17:c) shows the house's structure:

The main objective of this system is building an intelligent house that can be controlled through speech; for example, when the user says "eftah al bab" in the Arabic language (that's mean in the English language "open the door") the door is opened, also when he says "agloq al bab" in the Arabic language (that mean in the English language "close the door") the door is closed. Table 2 presents the components of the smart house in this system and how to manage and use. On the other hand, this system has two drawbacks that it is the user's far distance from the sensor and the presence of some acoustic noise in the place.
Table 2 Illustrate the smart component of house and how to manage and uses

<table>
<thead>
<tr>
<th>Component No.</th>
<th>Component Name</th>
<th>How To Manage</th>
<th>How To Use</th>
<th>The Accuracy Ratio Without sound Noise</th>
<th>The Accuracy Ratio Without sound Noise</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Door</td>
<td>By mic</td>
<td>When the user says &quot;eftah al bab&quot; the door is opened also when he says &quot;agloq al bab&quot; the door is closed.</td>
<td>100%</td>
<td>0%</td>
</tr>
<tr>
<td>2</td>
<td>TV</td>
<td>By mic</td>
<td>When the user says &quot;tashgeel al TV &quot; the TV is turned on, also when he says &quot;etfa' a al TV &quot; the TV is turned off.</td>
<td>100%</td>
<td>0%</td>
</tr>
<tr>
<td>3</td>
<td>Light</td>
<td>By mic</td>
<td>When the user says &quot;tashgeel al Dau'a&quot; the light is turned on also when he says &quot;etfa'a al Dau'a&quot; the light is turned off.</td>
<td>100%</td>
<td>0%</td>
</tr>
<tr>
<td>4</td>
<td>Fan</td>
<td>By mic</td>
<td>When the user says &quot;tashgeel al meruaha&quot; the fan is turned on also when he says &quot;etfa'a al meruaha&quot; the fan is turned off.</td>
<td>100%</td>
<td>0%</td>
</tr>
<tr>
<td>5</td>
<td>Gas Alarm</td>
<td>Sensor</td>
<td>The alarm devise is done automatically</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>6</td>
<td>Thieves Alarm</td>
<td>Sensor</td>
<td>The alarm devise is done automatically cally</td>
<td>100%</td>
<td>100%</td>
</tr>
</tbody>
</table>

Figure 17: The house's structure in different angles in (a), (b) and (c)

The comparison between previous studies that summarized in the introduction section and our proposed system is shown in Table 2.
### Table 2 Comparative scheme analysis for some datasets

<table>
<thead>
<tr>
<th>References</th>
<th>Year</th>
<th>Component of the smart house</th>
<th>The Beneficiary category</th>
<th>The smart component Managed By</th>
<th>language</th>
</tr>
</thead>
<tbody>
<tr>
<td>[7]</td>
<td>2007</td>
<td>Intelligent bed, intelligent wheelchair, and robotic hoist for effortless transferring user between bed and wheelchair</td>
<td>Elderly and people with disabilities</td>
<td>User's commands as well as to the recognized intentions of the user</td>
<td>-</td>
</tr>
<tr>
<td>[8]</td>
<td>2012</td>
<td>cupboards in the kitchen, fridge and bathroom cabinet, indoor doors, windows</td>
<td>elderly and frail population</td>
<td>Voice and tactile command</td>
<td>French</td>
</tr>
<tr>
<td>[9]</td>
<td>2019</td>
<td>Smart house with a living room, kitchen, bedroom and bathroom</td>
<td>All classes of people</td>
<td>Intent classification directly from raw speech inputs</td>
<td>French</td>
</tr>
<tr>
<td>[10]</td>
<td>2019</td>
<td>Device interface, a multi-modal interface, and a modal multi-modal dialog interaction engine. The multi-modal dialog interaction engine obtains input signals from the multi-modal interface (e.g. cell phone, browser, game console, sensors, etc.) where signals can represent different modalities of input (such as voice, images, gestures, touch, sensor data, and so on).</td>
<td>Receive input from users through sensing devices like a smart phone, a tablet computer or a laptop computer as voice, text, or from medical sensor and optical sensor. Users interact with system from inside the household or from remote locations</td>
<td>English</td>
<td></td>
</tr>
<tr>
<td>[12]</td>
<td>2020</td>
<td>Vacuum cleaner, a refrigerator and a television</td>
<td>All classes of people</td>
<td>Ontological engineering and multi-agent technology, voice control involves specific phrases and in a specific language</td>
<td>Specific phrases and in a specific language, which are embedded in each object</td>
</tr>
</tbody>
</table>

**Proposed Smart house System**

| Door, TV, Light, Fan, Gas Alarm, Thieves Alarm | Elderly and people with disabilities | Mic and Set of sensors | Arabic |

### Conclusion

In this paper, the major goal is to construct a smart home that is based on detecting and recognizing speech in the Arabic language to add more facilities to our lives and help some people (e.g. elderly or disabled people who live alone and have difficulty to move or use and handle the house components). To solve this problem, we built a smart house system that depends on speech detection and recognition. The system will distinguish speech and achieve its own function. For example, the door, the fan, the light, also the television in terms of dealing with them in the opening and closing via speech. Furthermore, the proposed system can sense fire and gas and alarm devise for detection of thieves which is done automatically. In future, we can add this effort to other components of the house, such as a refrigerator, cooker and the rest of the house furniture and execute...
it at the same principle. Moreover, we can execute this work in our homes where we live. On the other hand, the main disadvantage of this proposed system is that it doesn't work efficiently, if there is any noise in an environment.

Acknowledgements

The authors would like to thank the Mustansiriya Univ. (www.uomustansiriyah.edu.iq) Baghdad, Iraq, and University of Technology (https://uotechnology.edu.iq) Baghdad, Iraq, for supporting this work.

References


Proceedings of the First IEEE International Conference on Pervasive Computing and Communications (PerCom 2003), Fort Worth, TX, USA, 23–26 March 2003; 531–536.

Biographies of Authors

Kawther Thabet Saleh received the B.Sc. in Computer Science from Mustansiriya University - Iraq (2005), and M.Sc. in Computer Science from Al-Balqa’ Applied University- Jordan (2011). She has nearly 16 years of experience in teaching in Computer Science department, College of Education, Mustansiriya University. She has published over many papers in international and national journals and conferences. Her fields of interests include image processing, pattern recognition, Artificial Intelligence.
<table>
<thead>
<tr>
<th>Safa A. Ahmed</th>
<th>received her B.Sc. degree in Computer Science from University of Technology, Baghdad, Iraq, in 2005 and M.Sc. degree in Computer Science from Al-Balqa’ Applied University, Salt, Jordan, in 2011. Since 2006, she has been with the Department of Chemical Engineering at University of Technology, Baghdad, Iraq; where she is currently a full time Lecturer.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Haitham Salman Chyad</td>
<td>He was born in Baghdad, Iraq, in 1979. He received his M. Sc. Degree in computer system from Osmania University, India, Hyderabad, in 2016. He joined a programmer position in the laboratories of the Department of Computer Science at Al-Mustansiriya University for the period from 2007 to 2013 and manager of the division of electronic calculator of the College of Education at Al-Mustansiriya University and in addition to teaching in the Department of Computer Science.</td>
</tr>
</tbody>
</table>