Smart Homes with Voice Activated Systems for Disabled People

Bekir Busatlic [1], Nejdet Dogru [1], Isaac Lera [2], Enes Sukic [2]

[1] *International Burch University, Sarajevo, Bosnia And Herzegovina*
[2] *Department of Informatics, University of the Balearic Islands, Palma, Spain*

Abstract **– Smart home refers to the application of various technologies to semi-unsupervised home control It refers to systems that control temperature, lighting, door locks, windows and many other appliances. The aim of this study was to design a system that will use existing technology to showcase how it can benefit people with disabilities. This work uses only off-the-shelf products (smart home devices and controllers), speech recognition technology, open-source code libraries. The Voice Activated Smart Home application was developed to demonstrate online grocery shopping and home control using voice comments and tested by measuring its effectiveness in performing tasks as well as its efficiency in recognizing user speech input.**

Keywords **– Voice Activated Smart Home, Voice controller, helping people with disabilities.**

1. Introduction

According to the World Health Organization (WHO) more than one billion people in the world live with some form of disability, of whom nearly 200 million experience considerable difficulties in functioning. In the years ahead, disability will be an even greater concern because its prevalence is on the rise. This is due to ageing populations and the higher risk of disability in older people as well as the global increase in chronic health conditions such as diabetes, cardiovascular disease, and mental health disorders.

DOI: 10.18421/TEM61-15
https://dx.doi.org/10.18421/TEM61-15

Corresponding author: Nejdet Dogru,
International Burch University, Sarajevo,
Bosnia And Herzegovina
Email: nejdet.dogru@ibu.edu.ba

Smart houses can have a strong, positive and everyday impact on persons with physical disabilities and older persons, giving them privacy and comfort of their own home, yet allowing them to perform everyday tasks more conveniently and making them feel that they live an ordinary life, not confined to a hospital or in a special nursing home. This can potentially result in reduced the medical care costs per person, while at the same time improving their quality of life.

This application will help persons who have difficulty to move one point to another, such as very old senior citizens, patients with limited physical abilities. Users will be able to order their groceries from their houses so that they do not need to go to market frequently or they will be able to open the main door with voice comment for their family members or friends. Turning on and off lights, or any other devices will be possible without moving from their chair or bed. This system will decrease users' dependency on other people. Since application allows implementation of third party services, contacting family members or caretakers in case of emergency using voice comments can be easily implemented.

2. Related Work

A group of researchers [1] described the development of smart home environments for assisted living, with the key element being a voice user interface with additional capabilities such as gestures. They presented a universal platform for the control of smart home environments (mainly household appliances and assistive tools) oriented to the people with motoric disabilities. The key properties of the proposed platform were scalability and universality. The platform was composed of usual t and relatively cheap hardware elements, making it is easy to rescale it and involve a different number of controlled appliances and tools. The preferable mode of control chosen was voice based interface (VUI), which was realized using a hybrid recognizer adapted to identify the Lithuanian voice commands and a proprietary Hidden Markov Modeling (HMM) based Lithuanian speakerindependent recognizer to find the optimal decision. This approach enabled the minimization of the resources necessary to build the voice based interface. The achieved performance was high and acceptable for the vast majority of users involved in the tests.

The topic of smart home for elderly was evaluated in a study [2]. Authors presented the concept of smart home as an important goal of development with strong social and economic motivations. They explained how smart homes are a solution to many existing problems in the society, by making life more pleasant for aged/physically weak and/or disabled persons. They classified smart homes by design and installed technologies, specifying and commented on important components, their characteristics and requirements. In their review they tried to show that the development strategy of the intelligent home-installed technology has changed from the design of separate devices (at the beginning) to a form of integrated system arrangement where many home-installed devices communicate with each other and synchronously serve/monitor different parameters of the house. The first developments of home automation devices were simple hardware-oriented, while recent trends are oriented toward intelligence algorithms, where the software solution is the main part of the design. Home installed technology is further becoming oriented towards custom-tailored design, where modular components will meet individual user's needs. Authors suggested that in the future, some innovations in smart houses for persons with disabilities may be transferred to the nondisabled people, helping them with early diagnosis. In addition, they claim that wearable monitoring systems will become an important instrument for prolongation of the life expectancy in the near future.

Another study [3] is aimed at developing a system for vocal commands for home automation, which could be implemented in the personal home of elderly people or persons with disabilities. Significant challenges were identified that need to be overcome for the implementation of such solutions, particularly speech analysis in distant conditions, automatic speech recognition of elderly people or of expressive speech and more generally context aware interaction. Additionally, such solutions must be validated in real conditions with potentials users, which is why they conducted their evaluations in a smart home by potential users. They presented and evaluated a system for vocal command recognition and a system for speaker identification aiming at delivering this information to an intelligent controller in charge of home automation driving. Through the analysis of the results, authors derived some research issues due to real scenario conditions. The main lack was that people in the experiment, regularly deviated from the grammar (e.g., adding politeness terms or reformulation) and did not like the predefined chosen keyword. They also noted that in order to adapt the grammar to words that a user "naturally" uses in those situations, it would be better to have a learning algorithm, rather than imposing a grammar that had to be adjusted. Another issue was that their speaker recognition performance was very poor, with low precision when used by elderly people. In order to improve these lacks, one possibility would be to use information available from other sensors.. For example, if the speech recognizer system detects two speakers but the sensors only detect one person, the system can discover a state of inconsistency. Thus, the combination of different data sources can improve the automatic correction of the system.

3. Materials and methods

In this study we evaluate the efficiency and possibilities of using a voice activated smart home system as an aid for people with certain disabilities.

We implemente The Voice Activated Smart Home application (VASH) which is written in C# using Microsoft .NET framework. The application is the communication layer between the Speech Recognition Engine (Microsoft SAPI),Vera3 smart home gateway and other external services. We implemented an usage of online shopping on a grocery to demonstrate one service. However, since the application has a modular structure, any other service can be added easily.

3.1. VASH

The primary function of the VASH application is to perform the single use case it was designed for, receive user speech input, perform speech recognition, process the recognized text by performing the requested action and inform the user of the request result (i.e. success/failure, information requested) using speech synthesis. Delegates are also provided for GUI applications to know when SpeechRecognized and SpeechSynthesized events occur, so that they can be handled and the relevant information displayed on the GUI screen.

3.2. Speech Recognition Engine

There is a large variety of available software for speech recognition with no de-facto standards. The easiest implementation to use, as it is shown in article [4] is the Microsoft Speech Application Programming Interface (SAPI), which is not only built-in all current Windows operating systems, but also offers easily accessible managed code APIs [5].

3.3. Vera3 Controller

Vera3 is a smart controller that blends home control and Wi-Fi in a single box. It easily handles lots of devices, extends over the furthest range, and gives you the most flexibility to connect more gear. . Vera3 possesses a web GUI to operate over it. Vera3's capabilities allow it to be setup for IR control of entertainment systems and the ability to bridge with virtually any control platform (Z-Wave, Insteon, Serial, X-10, etc.)[6].

VeraDotNet

VeraDotNet is a .NET library for accessing capabilities of Vera home automation controllers. Vera controllers provide an VeraDotNet leverages the extensive http interface which provides access to many of Vera controller's functionality, to provide simpler access to Vera functionality from .NET capable devices. VeraDotNet encapsulates most of the complexity of interfacing with Vera3 controller and gives developer ability to make use of Visual Studio's IntelliSense functionality.

The VASH application implements two adapters: the SmartHomeAdapter for handling connection to the Vera smart home controller and the OnlineShopAdapter for handling Grandma Care web app communication.

The SmartHomeAdapter makes use of the VeraDotNet library to connect to the Vera controller and execute actions to devices.. The actions are called Luup requests that are Vera proprietary http requests for communicating with Vera controllers. The VeraDotNet library encapsulates those requests in managed code, so that is easy to use. An example of a Luup requests to control the switching of a light is shown in Fig. 1.

```
private bool SwitchLight(bool switchOn)
{
    try
    {
        //Get device by name
        var light = (BinaryLight1)miosEngine.GetDeviceByName("Light");

        if (light.SHaDevice1.CommFailure)
        {
            return false;
        }

        //Turn on or off
        light.SSwitchPower1.SetTarget(switchOn);
        return true;
    }
    catch (Exception exception)
    {
        return false;
    }
}
```

Figure 1. VeraDotNet library usage code sample

The OnlineShopAdapter connects to the Grandma Care web app through HTTP protocol. It uses the .NET HttpClient to communicate with the RESTful API. In Table 1, we show the RESTful services implemented in our case study.

Table 1. OnlineShop RESTful API

Method	Type	Parameters
api/Wishlist/AddToWishlist	POST	productId, quantity
api/Wishlist/RemoveFromWishlist	POST	wishlistItemId
api/Wishlist/GetWishlistItems	GET	n/a

Using VASH application

Once the application is running (Figure 2), it may be started using it by clicking on the large red square on the left side. The square will turn green and user may speak commands into the microphone.

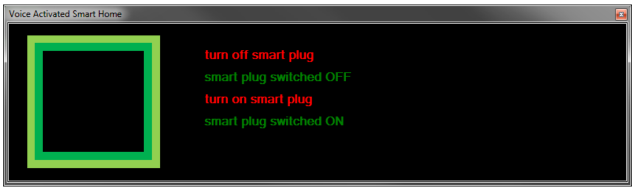

Figure 2. VASH application screenshot

The colors indicate the application status: red means the application is inactive, green means that it is waiting for user speech input, yellow means the application requires user input to complete action and blue means the application is processing user input. The text on the center-right shows the speech recognized and synthesized text, simply as visual guidance for testing the system.

User may use a number of commands, such as "Turn on the light" or "What is the temperature". If the application did not understand user command correctly or if it requires additional information to perform an action, user will hear a voice asking for confirmation or presenting user with a question (square will turn yellow), once he answer, the application will continue processing request. Once the application processes user request, it will tell the result or status of request (i.e. whether it succeeded or failed in performing the task).

In order to turn off the application, user simply click on the green square to turn off the speech recognition and click on the close button in the top-right side.

4. Results and discussion

The results of this work are evaluated according to several criteria to measure how effective the application was in achieving its target goals and with what efficiency. The criteria are effectiveness of the application (features implemented successfully), efficiency of the (speech recognition accuracy) and usage potential (market potential, future expandability).

In Table 2 the list of features found in the Vera3 controller are compared with the features implemented in the VASH application. Out of the seven major categories, the VASH application implements four features (main ones), two can be implemented with some additional work (camera and scenes) and only one (device configuration) is not suitable for implementation using the current application framework. The last feature cannot be easily implemented as it requires the speaker to configure a large number of devices and parameters, making it impractical to use.

Apart from the features implemented for the Vera3 controller, the VASH also implements several features for using the Grandma Care application, showcasing how it can be used to connect to any 3rd party web resources (RESTful API). This additional allows almost unlimited possibilities in expanding the application with more features, particularly with IoT devices, whether connected through a middleware controller or directly to the application.

Table 2. VASH application features comparison with Vera controller

Feature implementation	Vera controller	VASH application
Switch device on/off	YES	YES
Display sensor output	YES	YES
Control heating/cooling by setting temperature	YES	YES
Controlling door locks/unlock	YES	YES
Display camera image/stream	YES	NO Can be implemented with medium amount of work required.
Configure devices/rooms/scenes	YES	NO Configuration cannot be implemented with current application framework.
Control scenes	YES	NO Can be implemented with low amount of work required.

Speech Recognition Accuracy

To verify the accuracy of the Speech Recognition Engine used in this work, a test was conducted. The test consisted of 5 speakers (A, B, C, D, E), all having different English-speaking skills. Speakers A, B, C and D speak English with varying levels of proficiency: A has excellent, B has very good, C has average and D has a bad English speaking proficiency. Speaker E is a person who does not know English and is simply reading the commands written in English from a piece of paper.

Each speaker said 10 different commands and the results were assigned to five different cases (Table 3), depending on the outcomes.

Table 3. Speech recognition accuracy test case

Case	Case 1	Case 2	Case 3	Case 4	Case 5
Confidence rate	Above 0.9	Between 0.7 and 0.9	Between 0.7 and 0.9	Below 0.7	Above 0.9
Speech recognized correctly	YES	YES	NO	NO	NO
Result	Success	Success	Fail	Fail	Critical fail

Case 1: The application recognized the speaker's speech with a high confidence (above 0.9) and correctly recognized what was said. This is the best case scenario and signifies a very successful test result.

Case 2: The application recognized what the speaker said with a medium confidence (above 0.7, but below 0.9), but correctly recognized what was said. i.e. speaker answered with "yes" when asked whether the text that was recognized was the text that was originally spoken. This is another successful result, as the application understood correctly what was said, but with a medium confidence.

Case 3: The same as case two, except the speaker answered with "no" when asked by the application whether the text that was recognized was the text that was originally spoken. This is a failed test result, since the application failed to properly recognize what was said, but also indicated a medium confidence rate.

Case 4: The application's recognition confidence was low (below 0.7) and the application did not perform any action as a result, asking the speaker to speak again. This is a failed test result, but it is not critical, since the application reported a low confidence rate in what was said. This could also indicate the original speech was affected by a variety of factors (speaker proficiency, noise…etc.).

Case 5: The application recognized the speech with a high confidence (above 0.9), but the recognized text (and action performed) were incorrect. This indicates the worst case scenario and catastrophic test result.

A possible 6th case could also be identified as when the application recognized the speech with a low confidence rate (below 0.7), but correctly. This test case was not measured, since the application ignores all recognitions made with a low confidence rate. The probability of this outcome was also very low, meaning that it was of no importance to be measured.

Table 4. Speech recognition accuracy test

Case / Speaker	Case 1	Case 2	Case 3	Case 4	Case 5
A	10	0	0	0	0
B	10	0	0	0	0
C	8	2	0	0	0
D	8	2	0	0	0
E	7	1	1	1	0
Rate	43	5	1	1	0
Accuracy	86%	10%	2%	2%	0%
	96%		4%		

As seen from the results in Table 3, the overall accuracy of the system is 96%, with 12% of the tests requiring the user to confirm what was said, of which the majority was successful (over 80%). Only 4% of all tests resulted in a failure, with only 2% of all tests resulting in the application not being able to recognize what was said with a sufficient confidence. The tests did not result in any critical failures (case 5 result), indicating a false recognition with a high confidence rate.

5. Conclusion

In this study, the purpose and benefits of smart home technology was demonstrated, with a particular emphasis on persons with disabilities. The impact of disabilities was also researched and provided as evidence to support the beneficial effects of such voice activated smart home technologies.

The Voice Activated Smart Home application was developed and tested by measuring its effectiveness in performing tasks as well as its efficiency in recognizing user speech input. The effectiveness showed that the voice activated application can perform almost all tasks that are available to the Vera smart home controller, except the initial configuration. Some features were not implemented due to time and equipment restrictions, but given the code framework provisions were made so that implementation of those features was made possible. The efficiency of the system was measured to be very high, 96% when using, performing with less than 10-20% error rate in the worst case scenarios and with perfect 100% efficiency in the best case scenarios. This proves that the application is suitable for use in everyday scenarios.

The application's modular concept allows the expansion with additional features and the choice of technologies used, particularly in terms of capabilities by including more features, devices and technologies and security by including speaker recognition/identification technology. This study, through the showcase application that was developed, provides the foundations for further research and development in this field.

The VASH application can easily be expanded and improved upon with several key features such as speaker recognition and Windows 10 Cortana integration to allow for additional external services.

References

[1]. Rudzionis, V., Maskeliunas, R., & Driaunys, K. (2012, September). Voice controlled environment for the assistive tools and living space control. In FedCSIS (pp. 1075-1080).

[2]. Stefanov, D. H., Bien, Z., & Bang, W. C. (2004). The smart house for older persons and persons with physical disabilities: structure, technology arrangements, and perspectives. *IEEE transactions on neural systems and rehabilitation engineering*, *12*(2), 228-250.

[3]. Vacher, M., Lecouteux, B., Romero, J. S., Ajili, M., Portet, F., & Rossato, S. (2015, October). Speech and speaker recognition for home automation: Preliminary results. In *Speech Technology and Human-Computer Dialogue (SpeD), 2015 International Conference on* (pp. 1-10). IEEE.

[4]. Shi, H., & Maier, A. (2006). Speech-enabled windows application using Microsoft SAPI. *International Journal of Computer Science and Network Security*, *6*(9), 33-37.

[5]. Long, B. (2009). Speech Synthesis & Speech Recognition Using SAPI 5.1. Retrieved from blong.com: www.blong.com

[6]. Vera. (2016, 3 8). Vera3 Advanced Smart Home Controller. Retrieved from Vera: http://getvera.com/controllers/vera3/