Voice Activated Home Automation System

Sidhartha Chatterjee, Siddharth Shelton, Sanker Narayanan .S,
J. Hari Prasath, and Dr. B. Bharathi

Abstract- In recent years, Home Automation systems have seen rapid changes due to the introduction of various wireless technologies. Zigbee is targeted at applications that require low data rate, long battery life, and secure networking and has a maximum rate of 250 kbits/s. Well suited for periodic or intermittent data or a single signal transmission from a sensor or input device. This paper discusses the design of a system that allows the user to interact with it via his or her voice in natural language and maps user utterances to predefined instructions, including the switching of a light or electrical appliance. The system is intended to control all lights and electrical appliances in a home or office using natural language. Having a comfortable learning curve and being hands free by design, the system intends to make home control easier and more convenient.

Keywords—Home automation, Microcontroller, Voice Recognition, ZigBee, PocketSphinx

I. Introduction

The recognition and understanding of spontaneous unrehearsed speech remains an elusive goal in the area of computing. A human considers not only the specific information conveyed to the ear, but also the context in which the information is being discussed. For this reason, people can understand spoken language even when the speech signal is corrupted by noise. However, understanding the context of speech is, in turn, based on a broad knowledge of the world. This has thus been the source of difficulty and over forty years of research. The Home Automation industry is growing rapidly; this is fuelled by the need to provide supporting systems for the elderly and the disabled, especially those who live alone. Coupled with this, the world population is confirmed to be getting older. Home automation systems must comply with the household standards and convenience of usage. This paper details the overall design of a voice activated home automation system (VAHAS) which has been built and implemented. The automation centers on recognition of voice and uses low-power RF ZigBee wireless communication modules. The home automation system is intended to control all lights and electrical appliances in a home or office using voice. The system has been tested and verified. The tests involved a mix of 10 male and female subjects with different English accents. 5 different tasks were activated by each person.

II. Design

A. Hardware

1) Main Device: This device is responsible for practically all the computation and serves as the primary controller in the mesh network. The device is powered by the Raspberry Pi, which is a credit-card-sized Linux based microcomputer and is popular amongst the global DIY community. The Pi features the ARM1176JZF-S (ARMv6k) processor clocked at 700 MHz and features 512 MB of RAM. Capable of running Arch Linux ARM and requiring only 3.5 W of power for general functioning, it is a capable and efficient unit and serves dual purpose well. The Pi is connected to an audio interface via USB which in turn is connected to a small diaphragm condenser microphone. This sums up the audio input section of the device. In addition to the already mentioned peripherals, the Pi is connected via its GPIO pins to an XBee Series 2 XB24-Z7WIT-004 module from Digi which is essentially a serial ZigBee Network compatible RF module. This allows the device to communicate with the Switch.

2) Switch: The Switch is what lies at the other end of the network and its plurality has no limitations, since every appliance or power outlet to be controlled requires one of these to interface with the Main Device. The Switch is essentially powered by means of an 8-bit Atmel AVR microcontroller, pre-programmed with the Arduino boot loader that makes development easy using the Arduino integrated development environment (IDE) and C or C++ code. The microcontroller is interfaced via RS-232 with a

matching XBee Series 2 XB24- Z7WIT-004 module for RF communication with the Main Device. A solid state relay connected to the microcontroller enables it to switch high power circuits.

B. Software

1) PocketSphinx: Atop the Linux kernel running on the Main Device, is PocketSphinx, a lightweight speech recog- nition engine capable of keyword spotting and dictation using limited vocabularies and dictionaries tuned to the current application domain. Several changes have been made to the speech recognition code to make it suitable for real time input and to increase its accuracy in the tested environments. The output from the engine essentially consists of detected utterances which are then redirected to another component written using OpenDial.

2) OpenDial: OpenDial is a Java-based software toolkit to facilitate the development of robust and adaptive spoken dialogue systems. The toolkit itself is domain-independent and can be easily applied to any particular dialogue domain. Dialogue understanding, management and generation are ex- pressed in OpenDial through probabilistic rules encoded in a simple XML format. This is used to map user utterances to specific system actions, and keep the future possibilty of the system responding back to the user, open. Since this component is in Java, JNI is used to interface with the previously mentioned speech recognition component in C, and the output of this is sent back to the main funcion, which is responsible of sending the required instruction over RF to the particular Switch.

3) Firmata: Since low level pin access is required to ma- nipulate the appliance connected to the Switch in question, the Firmata protocol was adopted for communication between the Main Device and the Switch. Firmata is a generic protocol for communicating with micro-controllers from software on a host computer. It is intended to work with any host computer software package. There is a matching object in a number of languages.

III. Implementation

The following diagram illustrates the different algorithms involved in the system and their inter-working. As one can see, the speech from the user is captured by the microphone and is converted to digital information by means of the audio interface. The digital data is then transferred to the Pi via USB which is read constantly by the Listener code.

When the audio input crosses beyond a threshold of -10 dB, the algorithm calls the PocketSphinx code, which parses every utterance that includes a trailing and a leading 150 ms of silence. This inclues an inbuilt noise reduction algorithm that uses a CMM adaptive noise filter. The output from the PocketSphinx application is then sent as a parameter to the Java function that parses the input for intent.

After the intent is realised using the XML probabilistic rules file, the resulting instruction is sent via the XBee API to the Arduino Switch device in question, using the Firmata protocol.

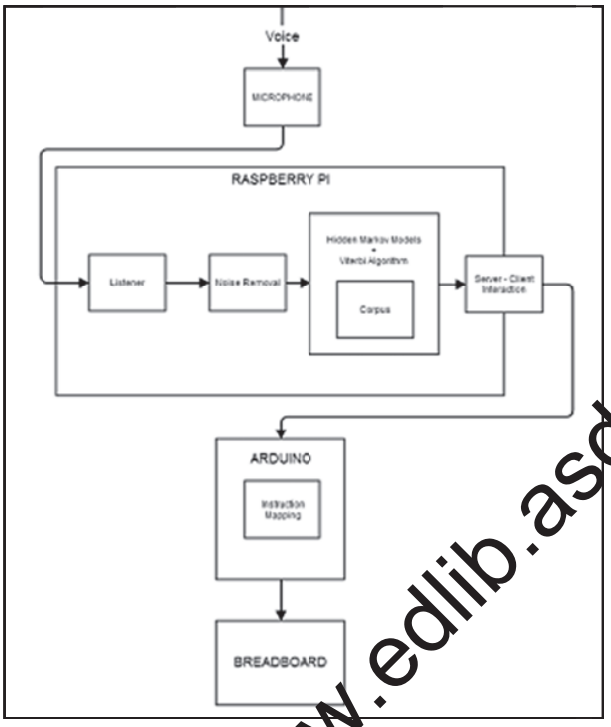

Figure 1. System Architecture

The Arduino which is constantly polling its serial input, parses the instruction and does the necessary action. This would usually involve writing 1 or 0 to a digital or an analog pin, which would result in the appropriate appliance switching on or off respectively.

IV. Results

The system was tested on a set of 30 males and females with different accents in English and each one of them was asked to execute five different instructions. The system was able to accurately identify the instructions 86.67% of the times. The system was also able to respond to a user's speech within an average of 150-200 ms, thus satisfying the requirements and resulting in an easy to use and responsive system.

V. Future Work

The voice recognition engine that is being used in the system is language independent and thus may be scaled to many other locales. Dictionaries and language models of other languages can be created using the tools available with the system, and the system can thus be trained to detect and parse other languages quite easily. The system could be made to run faster if deployed on better hardware. Since, most of the code is perfectly portable, scalability should not be a problem.

VI. conclusion

The VAHAS is a useful project for adults and physically disabled persons, who are not able to do various activities efficiently when they are at home and need ones assistant to perform those tasks. With the Voice Recognition along with ZigBee network we can eliminate the complication of wiring in case of wired automation and also it prevents to get up and down again and again to on/off appliances. ZigBee Home Automation provides operating range much higher as compared to Bluetooth and other wireless sensor

module. The system provides secure access to the home or office. Since it is language independent, the system can be scaled to many other locales. Also, since internet connectivity is not a requirement, the system may be used in locations without sufficient bandwidth penetration.

Acknowledgment

The authors would like to thank their friends and family for their constant support. They would also like to thank their advisor, Dr. B. Bharathi for her valuable inputs and her thoughtful suggestions during trying times.

References

1. A. Acero, Acoustical and Environmental Robustness in Automatic Speech Recognition, Boston, MA: Kluwer Academic Publishers, 1993.
2. J. Allen, Natural Language Understanding , 2nd ed. Menlo Park, CA: The Benjamin/Cummings Publishing Company, 1995.
3. G. Fant, Acoustic Theory of Speech Production,The Hague, NL: Mouton, 1970.
4. F. J. Owens, Signal Processing of Speech, New York, US: McGraw-Hill Inc., 1993.
5. R. Gadalla, Voice Recognition System for Massey University Smarthouse, Auckland, New Zealand: Massey University, 2006.