Interactive Voice-Controller Applied to Home Automation

Jinn-Kwei Guo, Chun-Lin Lu*, Ju-Yun Chang, Yi-Jing Li,
Ya-Chi Huang, Fu-Jiun Lu and Ching-Wen Hsu
Department of Computer and Communication, Kun Shan University, 710, Taiwan
**E-mail：cllu@mail.ksu.edu.tw*

Abstract

This paper describes the application of the embedded sound-controller, SUNPLUS SPCE061A, combined with Microchip 2.4GHz ZigBee modules, in Home Automation. Applying the speech recognition library provided by SUNPLUS, we use a multi-stage algorithm to identify which one of the target home devices should take the corresponding operation. The recognized control message is sent by the ZigBee module joined in the ZigBee network, routed and reached the home target device. Due to the multi-stage recognition phases, we can powerfully extend the capability of the original speech recognition library functions. Finally, some experimental results of an interactive voice-control lighting system demonstrate the facility, flexibility, scalability and feasibility of our work.

1. Introduction

For an easy life, there are many various CE (Consumer Electronics) home appliances being controlled and managed with different control units based on IR (Infra Red) [1]. On the other hand, Home Automation in ZigBee network is an emerging technique in recent years. Many home appliance manufactures are intended to control their products by ZigBee controllers. RF4CE Consortium was founded by Panasonic Corporation, Royal Philips Electronics, Samsung Electronics Co., Ltd. and Sony Corporation on June 12, 2008, to address increased demand for advanced functionality that is not currently available through infrared or other proprietary wireless technologies. RF4CE Consortium agreed to work with the ZigBee Alliance to jointly deliver a standardized specification for radio frequency-based remote controls on March 3, 2009. The ZigBee network is a candidate solution for home automation.

Many remote control systems based on ZigBee networks have been proposed in the literature. Wan-Ki Park et al. [1] developed a ZigBee based URC as the IR gateway of currently used home appliances. A remote lighting control system based on the CC2430 solution was proposed by Maoheng Sun et. Al [2]. Their controller is made up of CC2430, PCB Antenna, LCD panel and keyboard. The user can use the keyboard to operate the menu displayed on the LCD. Another PC based control software was designed to replace the hardware controller.

Ying-Wen Bai and Chi-Huang Hung [3] designed the remote power on/off with current measurement outlets for the power managed system. They used Visual Basic as the interface software for the design of the graphic user interface to provide a user-friendly operation of a typical home's outlets. The functions of the PC control software include the manual buttons, the reports of control statuses, and the real-time power consumption … etc.

Unlike [1] using ZigBee controller and ZigBee to IR conversion module, Jinsoo Han et al. [4] proposed an automatic standby power cut-off outlet as a ZigBee end device for the energy-saving room architecture. A ZigBee controller with IR code learning functionality was responsible for communicating with the power outlets and the dimming light. There are button switches on the controller manually operating the respective outlets and the dimming light. After the ZigBee controller learned the IR codes from a conventional IR controller, the user can press the button on the IR controller to wake up the respective appliance.

Generally there are many control units in our houses. The total number of the buttons on all the control units is more than one hundred. It is too complex to be operated well. The elders, children, and even disabled persons, may be confused by such a controller.

In this paper, we are interested in the replacement of the conventional control unit in the ZigBee network. We developed a hand-held interactive voice controller instead of the LCD and keyboard. The control scenario shown in Figure 1 is simple and easy. For example,

This paper is partly supported by NSC 97-3114-E-168 -001

when the user wants the light bulb in the living room to be switched on, the user just needs to say "living-room light on". The user speech is recognized and parsed by the proposed interactive voice-controller. The control message is then sent to the ZigBee network to operate the corresponding light.

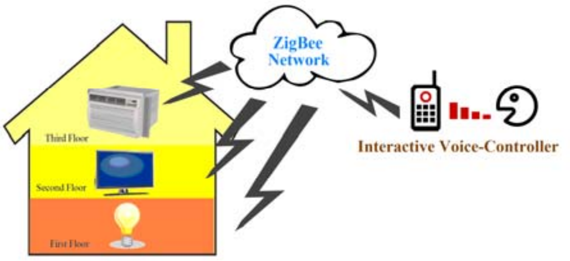

Figure 1. The application scenario

2. System Architectures

The proposed interactive voice control system architecture is shown in Figure 2. The main components of the interactive voice-controller are the ZigBee module and the sound controller for voice recognition. When the user speaks to the microphone, the sound controller recognizes the words. The recognized control message is then sent, routed by the ZigBee network, to the target ZigBee module. The corresponding ZigBee module will actuate the connected appliance via the connected driver circuit. The key hardware of the system is described as the follows.

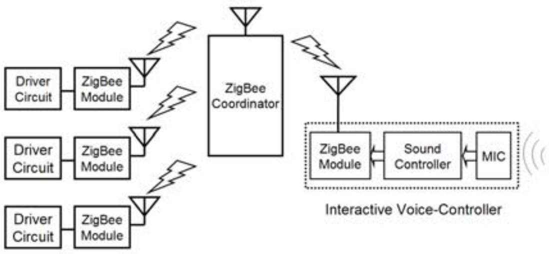

Figure 2. System architecture

2.1 Sound Controller

SUNPLUS SPCE061A is used for speech recognition in our interactive voice-controller. The sound controller has a build-in AGC microphone amplifier. The memory word is 16-bit wide. The 32k-word of flash memory is addressed from 0x8000 to 0xFFFF, and the 2k-word working SRAM is addressed from 0x0000 to 0x07FF.

We use the compact mini-61A module, as shown in Figure 3, in our experiments. SUNPLUS Company also provided a speech recognition software library called BSR library, which can recognize up to 5 speech words. The BSR voice model extracted 100 feature coefficients for each trained speech word and stored them into a SRAM buffer. As the result, there are 500 memory words in mini-61A to store these feature coefficients.

Figure 3. The voice recognition module mini-61A

2.2 ZigBee Module

The ZigBee module is a modification version of Microchip PICDEM Z 2.4GHz as shown in Figure 4. The layout of the main board is carefully designed to fit the space of the handheld unit. PIC18F4620 microcontroller is still used in the main board to take the advantage of easy migration of the ZigBee source code offered by Microchip. An 802.15.4 compliance 2.4 GHz transceiver module of Microchip is attached on the top of the main board to complete the ZigBee module.

Figure 4. The implemented ZigBee module

2.3 Driver Circuits

The driver circuit is necessary in the appliance end of the system. Comparing to the ZigBee module used in the handheld unit, a solid-state relay (SSR) and a power supply module are added as shown in Figure 5. The SSR is controlled by ZigBee module and then actuate the connected appliance. Totally three test sets are implemented in experiments as shown in Figure 6.

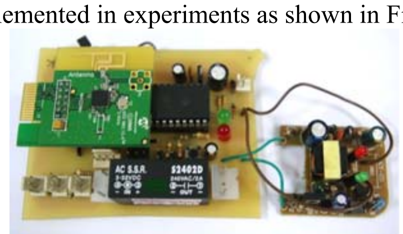

Figure 5. The driver circuit with ZigBee module

Figure 6. Three ZigBee light controller

3. Software Implementations

As mentioned in Section 1, the light bulb in the living room is switched on if the user says "living-room light on" to the voice-controller. The voice phrase consists of three parts, e.g, the location, the device and the operation. We define the command triplet as a command consists of three sub-commands. The first sub-command is the location command, such as the living room, the bedroom, … etc. The second is the device command, which is the target device to be controlled, such as the light, TV … etc. The last sub-command is the operation command, such as on, off, increase, decrease … etc. Followings are some legal examples of command triplets:

(Living-room, light, on)
(Bedroom, TV, volume up)
(All, light, off)

Due to the limitation of BSR library, only 5 speech words can be recognized. There are many location, device and operation sub-commands in our system. In order to extend the recognition capacity of the BSR library, we divided these words of sub-commands into several groups. Each group contains up to 5 words. Here we proposed two grouping methods in the following.

The first grouping method is the simple one. It follows the rule of ordinary command triplet. Hence we have three ordinary groups, e.g., the location group, the device group and the operation group. Additional groups may be followed to provide more operations.

The second grouping method may have more than three groups. The first group is the set of locations the same as method 1. The second group is the set of devices in the first location. The third group is the set of devices in the second location. The fourth group is the set of devices in the third location, and so on. Finally, the last group is the operation group.

The BSR library functions extracted 100 feature coefficients for each trained speech. A maximum of 500 words of memory were required to store feature coefficients for each group. There are only 32K words of flash memory in the sound controller. We allocated the flash memory addressed from 0x8000 to 0xDFFF for the execution code. The storage of BSR feature coefficients is started from 0xE000, and separated from 500 memory words for each group. Memory allocations in our implements are list in Table 1.

Table 1. Flash memory allocations.

Area	Address	Allocation
0	0x8000-0xDFFF	execution codes
1	0xE000	Gourp 1 coefficients
2	0xE1F4	Gourp 2 coefficients
3	0xE3E8	Gourp 3 coefficients
4	0xE5DC	Gourp 4 coefficients
5	0xE7D0	Gourp 5 coefficients
6	0xE9C4	Gourp 6 coefficients
7	0xEBB8	Gourp 7 coefficients

Initially all the flash memory data are cleared when the execution code is firstly programmed. At the first time start of the sound controller, the feature coefficients are empty in flash memory and the training algorithm is then executed. If the feature coefficients are found (e.g, the 2nd start), the training algorithm is skip to enter the recognition procedure.

3.1 Speech Training Procedure

The speech training procedure calls the BSR_Train() function in BSR library to get the speech feature coefficients. The 100 coefficients of a speech word are stored in a SRAM buffer (i.e., the BSR_SDModel[] array) after the speech word was trained. When 5 speech words in a group are already trained, the coefficient buffer is full. Then the 500 coefficients are saved to the flash memory allocated for that speech group. The BSR_ExportSDWord( ) function is called to export the coefficients.

A multi-stage training algorithm is used in the training procedure. The speech words in one group are trained in one stage. Speech words in groups are sequentially trained stage by stage. When the training process is complete in each stage, The coefficients are copied from the BSR_SDModel[]. After all the speech words in all groups are trained, the voice-controller can be applied to home automation.

3.2 Speech Recognition Procedure

The speech recognition procedure calls the BSR_GetResult() function to identify the speech words. Before the recognition process is started, the speech feature coefficients in a group must be stored in the SRAM buffer. The BSR_ImportSDWord() function is called to check if the BSR_SDModel[] is loaded successfully.

When the user talks to the voice-controller, the speech word is recognized. A multi-stage algorithm is used in the recognition procedure. The command triplet needs to be processed in multi-stages. For example, in grouping method 1, the feature coefficients of the location group are loaded into the SRAM buffer in the first stage. After the speech word of location sub-command is identified, the coefficients of the device group are loaded in the second stage. Finally the operation group coefficients are loaded in the third stage.

3.3 The Voice Interaction

The voice interaction is very important in complicate operation procedures. Proper hints and responses in voice will make the user feel easy in the operations. For the reason, the controller not only must offer a hint before the user giving an oral command, but also must answer the oral command perceptively. Hence a lots of voice library, such as “location name”, “please”, “living room”, “light”, “is recognized”, …, are built while the voice controller is programmed.

4. Experiments

Using grouping method 1, we have implemented 4 groups of speech words, as shown in Table 2. Group 1 is the location group, plus two additional commands. The “All” command means the operation applied to all the devices selected in Group 2. For example, the user could turn all the light off. The “back” command abort the current stage and back to the previous stage. Group 2, the device group, has 3 devices. Group 3 is the operation group. The “continue” command in Group 4 is designed to control the previously used target device. For example, after the “Bedroom Light On” is sent, the user can say “Off” instead of “Bedroom Light Off”. In effect, the “continuous” command is the same as “back”. Finally, the “end” command stops the current operation and back to stage 1.

Table 2. Implemented group commands.

Group 1	Group 2	Group 3	Group 4
Living room	TV	On	continue
Kitchen	Light	Off	end
Bedroom	Air conditioning	Increase	
All	back	Decrease	
back		back	

An example of the interactive scenario is implemented as the following where VC states for voice controller.

VC: “Location name, please.”
User: “Living room.”
VC: “Living room is recognized. Device name, please.”
User: “Light.”
VC: “Light is recognized. Operation name, please.”
User: “On.”
VC: “Living room light is on now. Continue or End?”
User: “Continue.”
VC: “Operation name, please.”
User: “Off.”
VC: “Living room light is off now. Continue or End?”
User: “End.”
VC: “Thank you.”

5. Conclusions

Based on the ZigBee network, we proposed a voice control method applied to home automation in this paper. The voice-controller composed of a sound controller and a ZigBee module plays the role of a user interface. Using the interactive voice-controller, the user can naturally and easily talk to the controller to control the home devices.

The limited recognition capability provided by SUNPLUS Company is extended by the multi-stage algorithm. This shows some scalability and feasibility of our algorithm.

References

[1] Wan-Ki Park, Chang-Sic Choi, Jinsoo Han and Intark Han, ”Design and Implementation of ZigBee based URC Applicable to Legacy Home Appliances”, IEEE International Symposium on Consumer Electronics (ISCE 2007), Irving, TX , June 20-23, 2007, pp.1-6.

[2] Maoheng Sun, Qian Liu and Min Jiang , “An implementation of remote lighting control system based on ZigBee technology and SoC solution”, International Conference on Audio, Language and Image Processing (ICALIP 2008), Shanghai, July 7-9, 2008, pp.629 – 632.

[3] Ying-Wen Bai and Chi-Huang Hung, “Remote power On/Off control and current measurement for home electric outlets based on a low-power embedded board and ZigBee communication”, International Symposium on Consumer Electronics (ISCE2008), Vilamoura, April 14-16, 2008, pp.1-4.

[4] Jinsoo Han, Haeryong Lee and Kwang-Roh Park, ” Remote-controllable and energy-saving room architecture based on ZigBee communication”, IEEE Transactions on Consumer Electronics, Vol.55, No.1, Feb. 2009, pp.264 – 268.