Optimizing the Functionality of a Voice Recognition System for Assistive Technology

Afeez Olalekan, Alex Page, Ying Sun, PhD
Department of Electrical, Computer and Biomedical Engineering
University of Rhode Island, Kingston, RI 02881-0805 USA

***Abstract*—The purpose of this study is to develop and optimize the functionality of a low-cost, light-weight system that allows for recognizing voice commands. The system utilizes state-of-the-art voice recognition electronics (VR Stamp, Sensory Inc.), which are programmed to recognize a small set of command words under the speaker independent mode. The functionality evaluated in this study includes issuing a nurse call signal and activating a single-switch controlled assistive device. The study also assesses the sensitivity and specificity of voice recognition by varying the speaking volume and the distance between the speaker and the microphone. The resulting system provides a reliable performance and is useful for voice activation of assistive devices to aid persons with disabilities.**

I. Introduction

The precise definition of voice recognition depends on the area of application. It usually means speaker recognition, but is sometimes referred to speech recognition as well. In this study, voice recognition is referred to the recognition of specific voice commands. The application area is the voice activation of a nurse call system or other assistive devices for persons with disabilities. Voice recognition technologies have been used for the integrated control of assistive devices [1] and the access of computers [2, 3]. These voice recognition technologies are based on software executed by a personal computer.

Previously we developed a voice-activated environmental control system [4], which was based on a voice recognition integrated circuit chip (HM2007). The system was successfully implemented with recognition accuracy about 85-90%, the occasional false activation was a major drawback. The accuracy of the system can be improved with improved voice recognition electronics and by limiting the vocabulary to a small set of command words. The purpose of this study is to develop a low-cost, light-weight system that can control assistive devices via voice commands without the need of a personal computer. The primary function of the system is to activate a contact switch for calling the nurse. It can also be used as a voice-activated switch to control other assistive devices.

II. Methods

A. Hardware

The system is based on the VR Stamp, the state-of-the-art voice recognition technology from Sensory, Inc. (Sata Clara, CA). As shown in Fig. 1, the core of the voice activation hardware is the VR Stamp module, which contains a complete

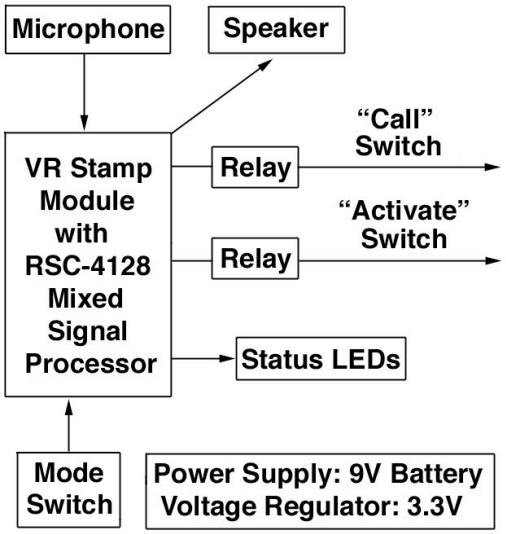

Fig. 1. Block diagram of the voice recognition system.

speech recognition system packed into a standard 40-pin DIP footprint (2" x 0.6"). The VR Stamp module includes the RSC-4128 mixed signal processor and 16-bit analog-to-digital and digital-to-analog converters, which is capable to hear (recognize voice commands) and to talk (generate synthesized speech). The VR Stamp evaluation board has been used to develop customized firmware in the C++ language. In addition to the VR Stamp, the circuitry also includes a speaker amplifier, relays for switch outputs, and status LED indicators. The system is powered by a 9V battery. A 3.3V voltage regulator is used to supply power to the VR Stamp unit.

The VR Stamp can be operated in either the speaker independent mode or the speaker dependent mode. For the speaker independent mode the voice commands to be recognized are from a standard library and preprogrammed into the processor. For the speaker dependent mode the voice commands are programmed for the specific speaker and can be re-recorded any time. For the present study, the focus is on the speaker independent mode, which does not require a training phase for the user and is universal for all users.

B. Software

The system is programmed to recognize two key words, "call" and "activate." The *call* command is intended to send a nurse call signal. To prevent unintentional issuing of the nurse call, the *call* command must be repeated twice within an 5-s

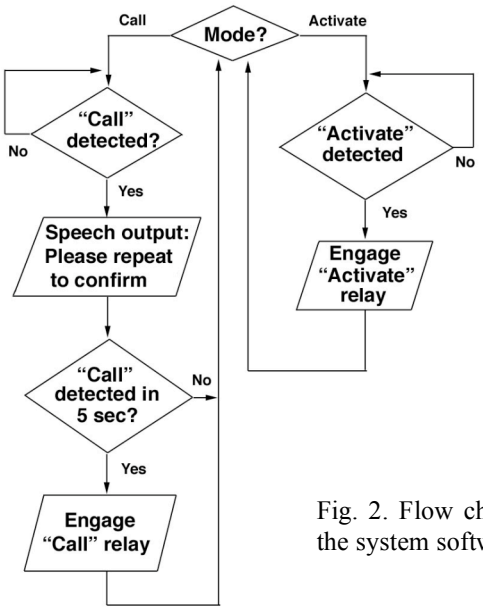

Fig. 2. Flow chart of the system software.

interval. After the first *call* command is recognized, the system outputs a voice message: "Please repeat to confirm." If the *call* command is detected within 5 s, the *call* switch signal is sent out by engaging the appropriate relay. The *activate* command is intended for controlling a single-switch operated assistive device. When the *activate* command is detected, the *activate* switch signal is sent out by engaging the appropriate relay. Figure 2 shows a flow chart of the software.

C. Evaluation of Sensitivity and Specificity

The performance of the system was evaluated by an experiment. To assess the sensitivity, the *activate* command was repeated 10 times and the number of successful detections recorded. To assess the specificity, the command "7-8" was used, because it sounds similar to "activate". The *7-8* command was repeated three times and the number of false detections recorded. This test was performed for three different distances between the speaker and the microphone, and for three different levels of volume (whisper, conversation, and loud).

III. Results

Table 1 summarizes the results of the performance test. The system was very sensitive (90%) to whispered commands within 1 foot of the microphone. Sensitivity reached 100% with louder commands at the same range. It still recognized approximately 50% of commands at a range of 5 feet, with a below-conversational volume, and about 80% of loud commands at that range. For the specificity test, a false detection was made when the system accepted the *7-8* command for the *activate* command. The system made several false detections except for two cases: whisper at 1-2 inches and at 5 feet. It was also observed that the system was less sensitive to the *call* command than the *activate* command. The system seemed to prefer a specific accent for *call* in order to work successfully.

Distance	Volume	True positive Using "activate"	False positive Using "7-8"
1-2 inches	whisper	9/10	0/3
	conversation	9/10	1/3
	loud	10/10	2/3
1 foot	whisper	9/10	1/3
	conversation	10/10	2/3
	loud	10/10	2/3
5 feet	whisper	5/10	0/3
	conversation	5/10	2/3
	loud	8/10	3/3

Table 1. Performance evaluation for the system with "activate" as the pre-programmed command.

IV. Duscussion

This project has resulted in a voice recognition system that is light-weight and low-cost (<$100). The system shows a good sensitivity. However, for commands that sound similar, such as *activate* vs. *seven-eight* in this study, the specificity can be relatively poor. This situation should be improved by utilizing the speaker dependent mode, whereby the system is trained to detect specific voice commands from a specific user. Future research will include a more extended study on the system performance by involving more speakers with different gender, age, and accent. To reduce the interference from the background noise, the use of a noise-cancellation microphone will also be investigated.

Acknowledgement

This study was supported in part by the URI Partnership in Physiological Measurements and Computing and was conducted in a special topic course taught by Dr. Robert B. Hill of the Biological Sciences Department at the University of Rhode Island. Students who contributed to an earlier prototype of the voice recognition system include Thomas Cavaliere, Shyla Booker, and Andrew Aubee.

References

[1] D. Ding, R.A. Cooper, B.A. Kaminski, J.R. Kanaly, A. Allegretti, E. Chaves, S. Hubbard. Integrated control and related technology of assistive devices. *Assist Technol.* 15(2): 89-97, 2003

[2] L.M. Boulton. Computer hardware and software to assist the visually impaired and blind. *Aust N Z J Ophthalmol.* 21(1): 7-14, 1993.

[3] Fried-Oken M. Voice recognition device as a computer interface for motor and speech impaired people. *Arch Phys Med Rehabil.* 66(10): 678-81, 1985.

[4] H. Jiang, Z. Han Z, P. Scucces, S. Robidoux, Y. Sun. Voice-activated environmental control system for persons with disabilities. *26th Northeast Bioengineering Conference:* 167-168, April 8-9, 2000.