

SPEECH RECOGNITION TECHNOLOGY FOR DISABILITIES EDUCATION

K. WENDY TANG

RIDHA KAMOUA

VICTOR SUTAN

OMER FAROOQ

GILBERT ENG

WEI CHERN CHU

GUOFENG HOU

Stony Brook University, New York

ABSTRACT

Speech recognition is an alternative to traditional methods of interacting with a computer, such as textual input through a keyboard. An effective system can replace or reduce the reliability on standard keyboard and mouse input. This can especially assist dyslexic students who have problems with character or word use and manipulation in a textual form; and students with physical disabilities that affect their data entry or ability to read, and therefore to check, what they have entered. In this article, we summarize the current state of available speech recognition technologies and describe a student project that integrates speech recognition technologies with Personal Digital Assistants to provide a cost-effective and portable health monitoring system for people with disabilities. We are hopeful that this project may inspire more student-led projects for disabilities education.

1. INTRODUCTION

Speech recognition allows “hands-free” control of various electronic devices. It is particularly advantageous to physically disabled persons. Speech recognition can be used in computers to control the system, launch applications, and create “print-ready” dictation and thus provide easy access to persons with hearing or

vision impairments. For example, a hearing impaired person can use a microphone to capture another's speech and then use speech recognition technologies to convert the speech to text. A vision impaired person, on the other hand, can use speech technologies to control electronic devices and launch applications.

Furthermore, voice recognition software could make a significant difference for many people with dyslexia. A learning disabilities researcher at Frostig Center in Pasadena, California, Marshall H. Raskind found that the technology can help make the overall college experience more rewarding and enjoyable for individuals with dyslexia [1].

Over the last decade, some very good voice recognition software has been developed for traditional desktops and laptop computers. A history of speech recognition and transcription software can be found in [2] and is not repeated here. More recently, there has been a surge of research and development activities in embedded speech technologies in which speech recognition technologies are deployed in toys, MP3 players, car navigation and entertainment systems, cell phones, and PDAs (Personal Digital Assistants). Efficient speech recognition engines that handle noise and variations in speech and faster, bigger, and cheaper processors and memory chips all helped to propel the migration of speech recognition technologies to smaller devices. It is also becoming obvious that when equipped with speech recognition technologies, PDAs can become user friendly to the hearing- and vision-impaired community by allowing users to launch programs, set up calendar appointments to compose e-mail and word documents, look up e-mail clients, and vocally input task information.

In this article, we first provide an overview of current commercial speech recognition software for desktops and embedded speech technologies for smaller devices in Section 2. Section 3 describes a senior design project that uses speech technologies to provide a cost-effective health-monitoring system for people with disabilities. A summary and conclusion is included in Section 4.

2. CURRENT SPEECH RECOGNITION SOFTWARES

In this section, we provide a summary and overview of current speech recognition technologies. Both commercially available software for desktops and embedded speech technologies are included.

Speech Recognition Software on Desktops and Laptops

In this section, we provide a brief summary of currently available speech recognition software on desktops and laptops. We consider the following softwares: 1) Dragon Naturally Speaking Preferred; 2) IBM ViaVoice 10.0; Keyston Speech Master 5; and 4) Microsoft Office XP.

Dragon Naturally Speaking Preferred 6.0

Adequate training for this package was found by reviewers to take around 20 minutes, or 10 minutes with follow-up practice. With NS, the user can “speak” into any open window, such as a Word 2002 document. The software also allows the user to “surf the Web hands-free”—in a Web page, if the user says the first few words of a link then the browser will go to the linked page. However, several reviewers indicated that the software requires a computer with a considerable amount of processing power to produce excellent results.

As of January 2003, the package was available for roughly \$210, with license deals available for multiple users. A search of various Websites indicated that versions of this package were the most frequently available speech recognition software in the universities.

IBM ViaVoice 10.0

Training for this system takes between 20 and 40 minutes (depending on reviewer). Version 10 is the most recent one, and according to some reviewers gives accuracy scores of over 96%, making it possibly the best current package in terms of accuracy. Speech input, still resulting in a high level of accuracy, is possible up to a rate of 160 words per minute. However, the software requires a considerable amount of processing power to work very accurately, and over 0.5 gigabyte of hard drive space. As of January 2003, the package was available for around \$140.

Keystone Speech Master 5

This package is a combination of Dragon Naturally Speaking and Keystone ScreenSpeaker. The added functionality provided by the latter provides various spelling aids and word recognition support, making it suitable for people with dyslexia. However, as of October 2002, the package costs \$590, presenting a significant barrier to its use for individual students.

Microsoft Office XP (built in)

Speech recognition is included within the latest version of the popular Office range of software. This allows the user to both enter commands, such as “open file,” and to dictate text straight into an application. Added functionality enables the translation of text among a small number of languages.

A small portion (four paragraphs) of this article was created using this system. This proved frustrating at first, as the author became very conscious (and distracted) by the process of entering content, as opposed to concentrating on the content itself. However, accuracy improved significantly from the first paragraph to the fourth.

Embedded Speech Recognition Software

In [3], a good summary of companies that develop embedded speech products is available. These include Advanced Recognition Technologies, Ltd. [4], Babel Technologies in Belgium [5], Conversay [6], Fonix Corporation [7], IBM Corporation [8], Scansoft Inc. [9], Sensory Inc. [10], Speech Works International Inc. [11], 20/20 Speech [12], and Voice Signal Technologies Inc. [13]. Readers interested in more details may refer to [3] and the companies' Web pages.

3. A SENIOR DESIGN PROJECT WITH SPEECH RECOGNITION

In the previous section, we provided a brief summary of current speech recognition software. In this section, we describe a senior design project using speech technologies with PDAs to provide a portable and cost-effective health monitoring system for physically challenged people. The remote health monitoring system was considered because people with health problems like Cardiovascular Disease (CVD) or other heart diseases need fast access to emergency services in order to increase their rate of survival. CVD is currently the leading cause of death in the United States. Therefore, constant monitoring is very important. The speech recognition assistant is integrated into the system to help interaction between the user and the system. This feature is especially useful for people with physical disabilities.

System Overview

The design emphasizes the safety of people with heart conditions, especially physically challenged people. When a person is experiencing heart problems, emergency 911 may not be accessible to the person at that particular instant, and the rate of survival of the person decreases with time. According to the American Heart Association [14], heart disease is one of the leading causes of death for both males and females in the United States. Figure 1 shows that heart disease or Cardiovascular Disease (CVD) is the leading cause of both male and female mortality in the United States, 49.24% and 55.10% respectively.

The solution is to try to reduce the time interval between dispatch and arrival of Emergency Medical Services (EMS). Thus, our goal is to design a device that can help reduce this time interval for patients, especially those that are physically challenged. After careful consideration, two features were considered: Health Monitoring System (HMS) and Speech Assistant.

Health Monitoring System (HMS)

This feature is designed to constantly monitor the condition of a user's heart. This system should consist of a heartbeat sensor, a processor, and wireless

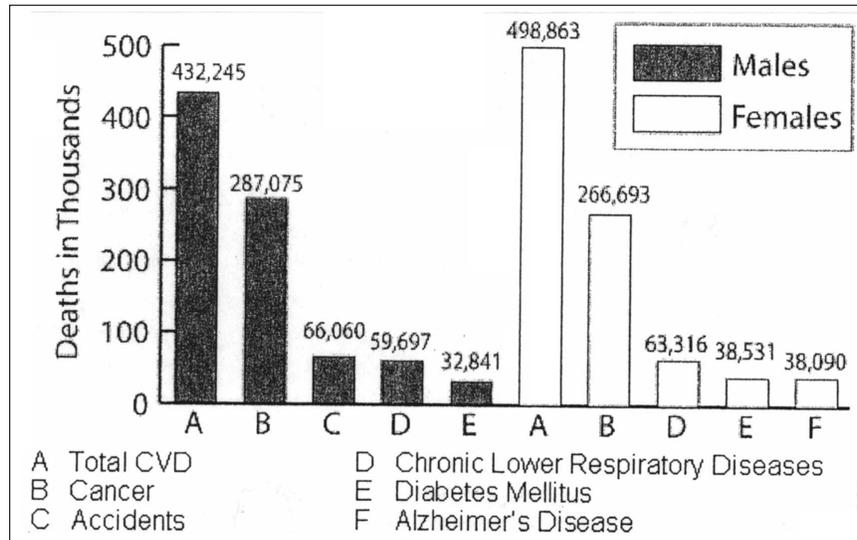


Figure 1. Leading causes of death of males and females in the United States.

communication support. The heartbeat sensor should be able to take an accurate measurement; the processor is needed to process the information sampled from the sensor; and the wireless communication is optional, but preferable for two reasons: 1) wireless communications make the system portable; and 2) current wireless technology provides many options, such as Bluetooth, or 802.11, and is relatively simple and cost-effective. The main task of the system is to notify authorized personnel when an emergency is detected. The notification can be in the form of e-mail, paging, or telephone messaging.

Speech Assistant

This feature is designed to assist the interaction between the user and the system. It is well worth considering because speech technology has improved greatly in the past decade. For portability, we chose to run the speech technology on a personal digital assistant (PDA). This feature should provide command launching capability and speech interaction between the user and the system. The speech interaction includes recognizing simple responses such as “yes,” “no,” and other common commands.

Following the design requirement, the final design will include a heart beat measurement system that communicates with a Personal Digital Assistant (PDA) with an integrated speech recognition assistant. There are remote health

monitoring systems available on the market, but none of them uses the capability of the PDA. Most devices are optimized to carry out only a certain task such as measuring heart beats or blood pressure. Our design, on the other hand, can be used as a remote monitoring processor with other additional capabilities such as word processing or a scheduler.

Some benefits of our design are as follows:

- Remote heart monitoring on people with a heart condition;
- Assisting the user in interacting with the system;
- Providing two-way communication between the user and emergency personnel (such as the patient's doctor);
- With future expansion, providing for an automatic external defibrillator; and
- With future expansion, allowing for other kinds of health monitoring and forming the precedent for using the PDA in the medical field (such as the SpO₂ level, blood pressure or even sugar level for diabetics).

Implementation and Engineering Consideration

The design was implemented by two teams: the hardware team and the software team. The hardware team's tasks were to design and implement the HMS module. The software team's tasks were to learn Windows CE programming and Microsoft Speech API. Toward the end, the two teams worked together to integrate both modules.

Hardware Team

The implementation of the hardware module is divided into several phases: research and design, implementation, and testing and debugging.

Research and design—The period of the research was approximately three months before the actual design process. The research covered several issues such as sensors, processors, and wireless communication.

- *Sensors*: The infrared sensor was used based on the fact that density in human blood can be measured with a simple circuit built from an infrared LED and a photodiode (see Figure 2). The data sampled is based on the amount of infrared light that is detected by the photoresistor. When blood is pumped into the capillaries of the finger, a slight change in density causes the change of light. This sensor is simple and accurate enough to measure human pulse. However, in certain people it may also detect the backbeat and throw off the accuracy of the reading of the actual heart rate. This technology was one of the earliest devices to detect a heart beat. Today's medical facilities use electrocardiograms (EKGs). EKGs are far more expensive and consequently, the infrared/photoresistor technology was chosen for this design.

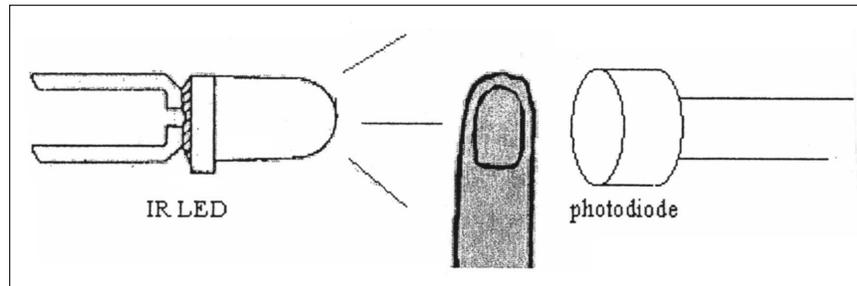


Figure 2. Infrared heart pulse sensor.

- *Microcontroller Board or Field Programmable Gate Array (FPGA):* A microcontroller board is targeted for larger design and developmental purposes, whereas FPGA is targeted for a smaller design. A microcontroller board provides the user with many features such as a serial port. However, many of its features are redundant for a smaller design such as the watchdog timer. Moreover, an FPGA takes less space than a microcontroller board, is inexpensive, and consumes less power. However, this project is still in its developmental phase. Thus, a PC104 microcontroller board was chosen. The PC-104 provides an RS-232 serial connection, supports ASM186, and provides memory expansion slot.
- *Serial or Bluetooth:* A serial and Bluetooth medium was considered in the design. A serial connection is simple, and is supported by PC104 microcontroller board, but it is not very convenient for the user, unless he or she is willing to wear a wire. As a result, a Bluetooth connection was used. The Bluetooth module is connected via serial to the PC104 board. It will be able to transmit data wirelessly.

Implementation:—Implementation of the hardware design took approximately one and one-half months. All ICs are mounted into an evaluation board. Power came from an old AT IBM case. The infrared LED and photoresistor were also mounted onto the evaluation board. Figure 3 shows the schematics for pulse sampling.

Testing and debugging—Testing and debugging are achieved by using Paradigm C++ that runs under the Windows XP environment. This phase took about two to three months before we achieved an accurate measurement of the heart beats. The following summarizes the problems encountered and the solution we provided:

- The first sensor was built into the evaluation board. The sensor gave out a reading with around 40% error, which is not tolerable at all. The error was

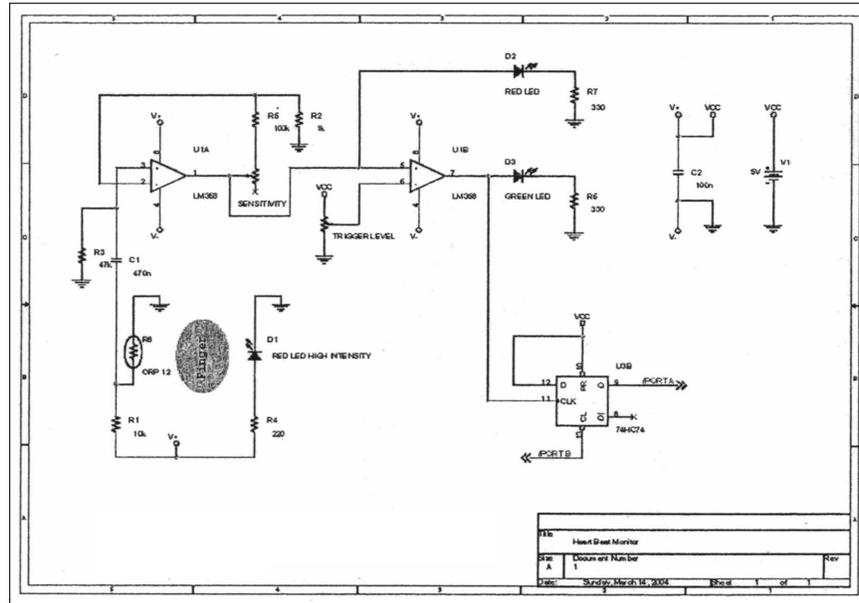


Figure 3. Infrared sensor.

assumed to come from the light dispersion and absorption. After several experiments, the problem was solved by building a black shield to block the light absorption from the air. The error was actually reduced to 25%, which is still not acceptable.

- The second design is to create a clip with the sensor. The sensor was mounted into the clip, and the clip was covered with black tape to prevent light dispersion. This solution reduced the error to less than 1%.
- The Bluetooth module was tried with windows hyperterminal and Bluetooth dongle (BluePC) for PC. This module is able to send the information to the PC without any problem.

Software Team

The software designs cover three phases: research and design, implementation, and testing and debugging.

Research and design—The software team worked on, and continues to refine, a Graphical User Interface program and a speech application. Based on the requirement of the design, there are several issues to be considered, as follows:

- *Remote Monitoring Program:* This program is written under Windows CE programming because it is a PDA based-system. The system is required to receive the pulse from the HMS through Bluetooth connection. The program then displays the information on GUI window. The program also provides a security feature, such as password-protected range modifier. The program also monitors the user's information, and sends an emergency message to emergency personnel if the user's pulse falls into an unsafe zone. Before an emergency message is sent, the program prompts the user for an input to confirm the emergency. The input can be in the form of stylus input, speech, or keyboard. The program assumes it is an emergency if there is no response from the user after 45 seconds.
- *Speech Application Assistant:* The initial decision on the speech application was to use commercially available software. Some examples are Scansoft's Dragon Naturally Speaking and Microsoft Voice Command. They provide capability to launch command, manual speech input, and Custom Command. However, the features in that software are fixed, and unnecessary for this project. The second option was Microsoft Speech Application SDK Beta 4. This SDK is targeted for Web application only, and not suitable for Windows or console application. The final decision was to use Microsoft Speech Application Program Interface (SAPI) 5.0 for Windows CE. The SAPI is free and Pocket PC compatible, but a third-party speech recognition engine is needed. The tasks for the speech program are as follows:
 - prompt the user for speech input
 - interact with the user

Implementation:—The software implementation took approximately two to three months. The Windows CE Alert System was written to communicate two ways with the Speech Recognition Assistant. When the Windows CE Alert System wants to confirm an emergency, it transfers the control to the Speech Recognition Assistant. The Speech Recognition Assistant will transfer the control back to the Windows CE Alert System after 45 seconds, or speech input is received. Figure 4 illustrates the software implementation.

Testing and debugging—This phase took approximately a four-month period and was started concurrently with the hardware testing and debugging.

- Windows CE GUI programming was tested on Pocket PC 2002 Platform SDK emulator. The most common error was a linking error. It was solved by linking the right library and using the right platform. The GUI Program was successfully deployed into the PDA.
- Speech Application Assistant was tested on Microsoft Visual Studio.NET 2003 Professional. The first experiment was to run the speech application on an emulator in a desktop before deploying it to the PDA. The Speech application was successfully tested on the emulator. The sample program was

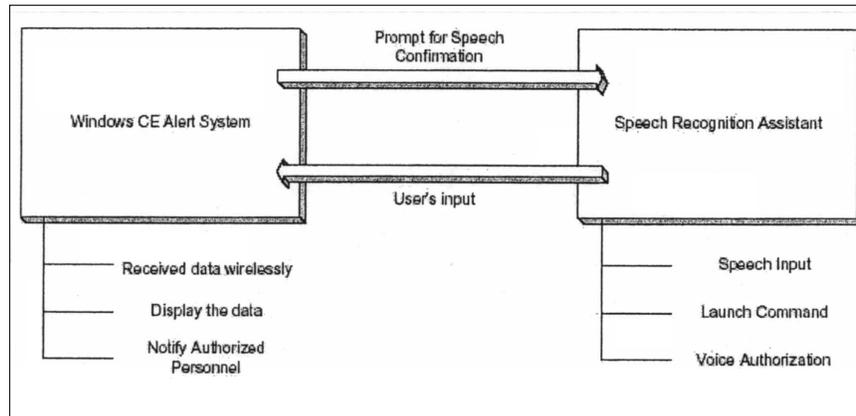


Figure 4. Software implementation.

able to receive speech input and display the input. The second experiment was to run the speech application on PDA. However, we had some problems deploying the program into the PDA. We are currently working with Microsoft technical support to solve this problem. We are hopeful that a solution can be identified in the near future.

Tools

The following is a summary of various tools in our design:

1. **Paradigm C++:** Paradigm C++ is an embedded system development software by Paradigm System, Inc. The software is developed for assembly, linking, locating, and debugging assembly language and C/C++ code for embedded microprocessor design. Paradigm C++ was chosen based on the fact that the hardware team has 1.5 years experience working with it.
2. **ORCAD Capture 9:** ORCAD Capture is a simulation and drawing tools from Cadence. For this project, it was used for a schematic purpose only.
3. **Microsoft Visio 2003 Professional:** Microsoft Visio 2003 professional is software developed for planning; for example, block diagram and flowchart. Visio was used to draw the flowchart of the design.
4. **eMbedded Visual C++ 3.0:** eMbedded Visual C++ is a software development tool targeted for Pocket PC 2002. This software has an advantage over Visual Studio 6 in developing C++ program because it can be used to emulate programs on the platform SDK. This software is used to write the Windows CE GUI.

5. Microsoft Visual Studio.NET 2003 Professional: MSVS.NET 2003 Pro is needed to develop software with SASDK compliant. This software also provides an environment for Speech API compliant application. This software was used to test the speech assistant.
6. Hyperterminal: Hyperterminal is a Windows application that supports port connection simulation. Hyperterminal was used to test BlueCOM on a PC. The procedure was done by sending a character of data, even parity, 8 bits with 19200 baud rate.
7. BluePC: BluePC is a Bluetooth dongle that acts as a virtual serial port for PC, as well as other ports. It was used to setup a virtual COM port. Hyperterminal was then used to connect to the COM port and confirmed the data sent from the BlueCOM.

4. SUMMARY AND CONCLUSIONS

Speech recognition is an alternative to traditional methods of interacting with a computer, such as textual input through a keyboard. An effective system can replace or reduce the reliability on standard keyboard and mouse input, assisting in particular dyslexic students who have problems with character or word use and manipulation in a textual form, and students with physical disabilities that affect their data entry or ability to read and therefore to check what they have entered.

In this article, we first provide a summary of commercial speech recognition software and then describe a senior design project that provide a cost-effective and portable health monitoring system with speech assistants to help people with disabilities.

The project has two main features: a health monitoring system and a speech assistant. The Health Monitoring System constantly monitors the user's heart, while the Speech Assistant helps the user to interact with the system. Such constant monitoring feature will reduce the 911 access time and increase the rate of survival of the user. The Speech Assistant helps the user with input to the PDA. This feature is very important in making PDAs usable to physically challenged people.

A working prototype has been successfully created. The current status of the prototype is that we can provide an accurate measure of the heart beats and wireless communications are achieved via Bluetooth. However, we are still having some difficulty with the Speech Assistant. We have created and successfully run the speech coding on an emulator. However, we have had some problems deploying the program into the PDA. We are currently working with Microsoft technical support to solve this problem.

There are a few remote health monitoring systems available on the market but none of them utilizes the capability of the PDA. Furthermore, most devices are optimized to carry out a single task only. The unique contribution of this

design project is that we have integrated the PDA with wireless communication and speech technology in our prototype. Besides heart beat monitoring, it is capable of assisting users with voice commands, providing two-way communications between the users and emergency personnel, and can be expanded to include monitoring of multiple factors, such as blood pressure or even sugar level for diabetics.

REFERENCES

1. *History of Speech Recognition and Transcription Software*, www.dragon-medical-transcription.com/historyspeechrecognition.html.
2. *Voice Recognition Software Benefits Dyslexics*, www.boiceability.com/Benefits_Dyslexics.htm.
3. J. Kumagai, *Talk to the Machine*, IEEE Spectrum, pp. 60-64, September 2002.
4. Advanced Recognition Technologies Ltd., www.artcomp.com.
5. Babel Technologies, www.babeltech.be.
6. Conversational Computing Corporation, www.conversa.com.
7. Fonix Corporation, <http://www.fonix.com>.
8. IBM Corporation, www.ibm.com.
9. Scansoft Inc., www.scansoft.com.
10. Sensory Inc., www.sensoryin.com.
11. Speechworks International Inc., www.tmaa.com/tts/SpeechWorks_profile.htm.
12. 20/20 Speech, www.2020speech.co.uk.
13. Voice Signal Technologies Inc., www.voicesignal.com.
14. American Heart Association, www.americanhearts.org.

Direct reprint requests to:

Dr. K. Wendy Tang
Dept. of Electrical and Computer Engineering
Stony Brook University
231 Life Engineering Lab
Stony Brook, NY 11794-2350
e-mail: wtang@ece.sunysb.edu

Copyright of Journal of Educational Technology Systems is the property of Baywood Publishing Company, Inc. and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.