Speech Interaction with Handheld Computers in the Project54 System

Kristе Krstovski, Andrew L. Kun, W. Thomas Miller III
University of New Hampshire
Electrical and Computer Engineering Department
Durham, NH 03824 USA
{kristek, andrew.kun, tom.miller}@unh.edu

ABSTRACT

The Project54 system integrates electronic devices in police cruisers. This integration allows officers to interact with the individual devices using speech. The in-car Project54 system is extended by using distributed software components running on Windows-based handheld computers. These components allow the police officer to control all the in-car electronic devices from the palm of his or her hand. We present a speech user interface for the components running on the handheld computer and discuss how it can be used in the field.

INTRODUCTION

Today's police cruisers are equipped with a large number of electronic devices (lights, sirens, digital radios, scanners, radars, mobile computers, GPS units, barcode readers, video recorders, etc). In the Project54 system, these devices are controlled by integrated software components [1] running on an embedded computer and distributed components deployed on a handheld device [2]. Inside the cruiser, the officer can interact with the system using spoken commands and receives feedback from the system via synthesized speech. This allows officers to keep their eyes on the road and their hands on the wheel while using electronic devices.

Officers often leave their cars during traffic stops, accident investigations and emergency situations. In these situations, the officer could stay connected with the in-car devices through Project54 software running on a handheld computer. The most common mode of interaction with today's handheld computers is through a Graphical User Interface (GUI) - users are typically required to look at the GUI and tap the screen of the handheld in order to accomplish actions. However, in many cases officers need to keep their eyes on their environment in order to be able to act in potentially dangerous situations and they also need to keep at least one hand free in order to be able to respond to any physical attack. Thus, for officers, using one hand to hold the handheld computer, looking at a GUI and using the other hand to tap on the screen with a stylus is often impractical. In order to solve this problem, we created a Speech User Interface (SUI) for the handheld computer. Having a SUI allows the officers to keep their eyes on the environment and it also allows them to keep at least one hand free.

SYSTEM SOFTWARE ARCHITECTURE

We created the SUI by implementing a system software architecture that supports speech interaction. The block diagram of the Project54 system software architecture implemented on a Pocket PC is shown in Figure 1. At the center of the system is the Application Manager. The Application Manager implements the message coordinator object which provides a means for receiving messages from individual applications. The applications implement a message handler object which provides a means for receiving messages from the Application Manager. The Application Manager handles messages sent between applications by receiving messages via the message coordinator object and sending messages via the message handler objects, as shown in Figure 1. This architecture mirrors the architecture implemented on an embedded computer for the in-car Project54 system [2].

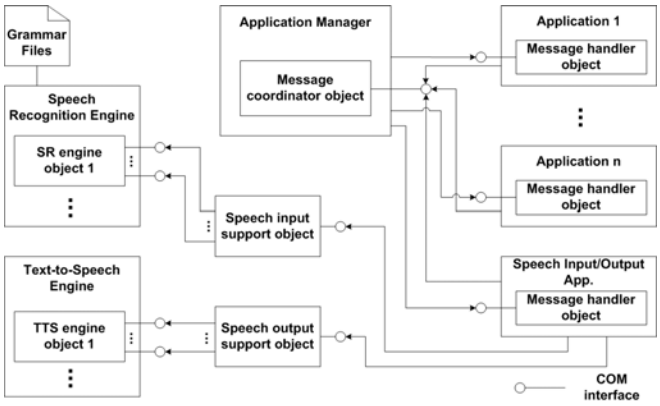

Figure 1. – The Project54 system software architecture allows interfacing with SR and TTS engines implemented on a Pocket PC.

Most of the applications running on the handheld computer control electronic devices. They are connected to the in-car devices through a wireless link [2]. The applications can be speech enabled. The applications do not interact with Speech Recognition (SR) and Text-To-Speech (TTS) engines directly. Instead they utilize a special application, the Speech Input/Output Application (SIOA). The SIOA receives recognition results from the SR engine via the Speech input support object and forwards these results to the appropriate applications in the form of text messages. Likewise, applications send text messages to the SIOA to

request speech output. The SIOA forwards these requests to the TTS engine for output via the Speech output support object. Each support object exports a set of interfaces. These interfaces are standardized for the Project54 system. Consequently, in order to use a new SR or TTS engine only the appropriate support COM object needs to change. The applications (including the SIOA), as well as the Application Manager, do not need modification.

WILL SPEECH INTERACTION WITH A HANDHELD COMPUTER WORK IN THE FIELD?

In order to improve speech recognition accuracy today's speech recognizers use grammars and press-to-talk (PTT) buttons. A grammar lets the recognizer know what rules user utterances will follow. For example a simple grammar may state that the user will only say one of two words: "male" and "female." This grammar may be useful when "telling" an application someone's gender. A PTT button lets the recognizer know when to "listen" to speech input. This reduces the probability that the SR engine will have to process sounds not directed at it (and potentially make a mistake).

We created a Speech input support object for a speech recognition engine developed for handheld computers by HandHeld Speech [3]. We also created a SUI Test Application, a speech-enabled application that lets us test basic SUI functionality. A screenshot of this application is shown in Figure 2.

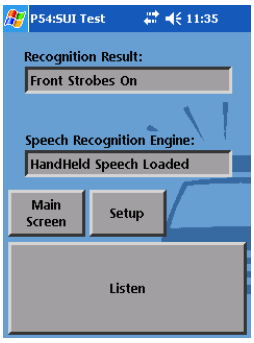

Figure 2. – The SUI Test Application.

The SUI Test Application implements a PTT button in the GUI – it is the large button at the bottom of the screen. While this button is pressed sound input from the microphone is fed to the recognizer for processing. The test application can be set up so that speech interaction is logically divided into interactions with different topics and corresponding grammars. Since the HandHeld Speech recognizer is speaker-dependent users have to first train the recognizer. Our preliminary tests with the SUI Test Application confirmed that the SUI is operational. Will speech interaction with the handheld be a viable approach to interacting with the handheld computer in the field? There are several issues related to this question. The first one is the quality of the speech signal that one can obtain with the handheld computer's microphone (internal or external). In the in-car Project54 system the quality of the speech signal is preserved by using a directional microphone. Our current setup uses the handheld computer's built-in microphone but, using the Test Application, we will experiment with external microphones as well. Another issue is that of organizing the user interface. In the in-car system, user interactions, and the necessary grammars, are organized around interactions with particular devices. We will need to determine what officers in the field will use the handheld computer's speech interface for, and run user studies grouping the necessary phrases into grammars in different ways. This will be done by speech-enabling multiple applications on the handheld computer. Finally, an important issue not related to the SUI design will be the ruggedness of the handheld device – if the handheld is to be deployed in the field it will have to withstand rough handling.

RELATED WORK

Microsoft's Dr.Who research project aims to eliminate the need for tiny styluses or minuscule keyboards by introducing the MiPad speech-enabled handheld computer [4]. Researchers at the IBM TJ Watson Research Center have introduced the Personal Speech Assistant which employs an embedded speech engine and uses speech as the primary mode of interaction [5]. While these solutions offer speech enabled interaction with the PDA they do not offer a way of using third-party speech recognizers.

CONCLUSION

We implemented the software infrastructure of a SUI for the components of the Project54 software running on a handheld computer. Currently different approaches to the SUI design are being tested in a laboratory environment. By the end of 2004, our tests will involve police officer subjects using the system for their everyday work.

REFERENCES

1. Andrew L. Kun, W. Thomas Miller III, William H. Lenharth, "Project54: Standardizing electronic device integration in police cruisers", IEEE Intelligent Systems, Volume 18, Issue 5, Pages: 10- 13, 2003.

2. Albert Pelhe, Nevenka Kozomora, Andrew L. Kun, W. Thomas Miller III, "Distributed components in the Project54 system", Workshop on Distributed Objects Research, Experiences & Applications (DOREA '04), Cancun, Mexico, 2004.

3. Gregory Gadbois, HandHeld Speech, LLC website, www.handheldspeech.com [Accessed: 7 June 2004]

4. X. Huang et al., "MIPAD: A Multimodal Interaction Prototype", ICASSP, Salt Lake City, UT, 2001.

5. L. Comerford, D. Frank, P. Gopalakrishnan, R. Gopinath, J. Sedivy, "The IBM Personal Speech Assistant", ICASSP, Salt Lake City, UT, 2001.