Issues in Delivering Mobile IT Systems to Field Users

J. H. Garrett, Jr. and J. Sunkpho
Department of Civil and Environmental Engineering
Carnegie Mellon University, Pittsburgh, PA, USA

1.0 INTRODUCTION

IT support for field personnel is starting to appear on bridge inspection sites, construction sites, and on the manufacturing floor. ENR magazine now regularly advertises palm computing-based systems for recording data on a construction site. This support can, and should, take on many forms and functions and needs to be carefully designed and evaluated like any other important tool used in the field. With careful design and implementation, IT devices for the field should have tremendous impact on the productivity of field personnel. If the devices are not carefully designed to take account of the field context and the abilities and preferences of the field personnel, field workers will not use these tools.

1.1 A Motivating Scenario

Imagine that a team of bridge inspectors is inspecting a large highway bridge. Some inspectors are visually inspecting the more accessible parts of the structure. They rapidly record what they are seeing using a speech interface to their data collection device. They compare the current condition to that reported during previous inspections by viewing any of the previous inspections reports stored in a central database. Some inspectors are querying sensor systems embedded in, or attached to, specific locations on the structure using a mobile device for interfacing with these various sensors. The inspectors are able to access the loading conditions and corresponding structural responses, and any changes in local material properties, of the structure since the last inspection period. Other inspectors are operating mobile inspection devices from a remote position somewhere on the bridge. They are able to move these devices to different locations so as to view and sense the condition of the underside of the bridge superstructure. They are also able to control and interpret the information being provided by these devices. Finally, other inspectors are using advanced condition assessment technologies, such as video imaging, infrared thermography and ground-penetrating radar to inspect certain elements of the structure, such as the bridge deck. As these inspectors are collecting and viewing the results collected from these various sensor systems, they are evaluating the conditions of the structure and the need for additional tests, with the help of decision support systems.

Each inspector is providing input directly into an inspection report being compiled as they perform different aspects of the inspection. In other words, they are all using some form of mobile IT device to efficiently and effectively support their specific field-oriented tasks. Nobody is sitting at a desk; they are on the bridge and moving about, climbing, and conducting their inspection tasks and visually confirming what the various sensors and condition assessment technologies are reporting.

In addition to being easy to use, the mobile IT systems provide intelligent, supportive functionality. For example, they help locate possible damage sites and assess their states of damage. When a defect or error is found, the location, nature and extent of the damage must be accurately described by the inspector. An intelligent inspection assistant presents the inspector with a detailed, context-specific description of the procedure to follow in assessing the damage once it has been discovered. For example, the system advises the inspector about where and how many times the rust and scale should be scraped from the girders on a superstructure so as to measure the remaining section. Knowledge Discovery in Data (KDD) techniques are used to help inspectors identify common damage locations on the bridge being inspected based on the data collected on other bridges in the database. Based on the locations and extent of damage on other similar bridges, under similar loading conditions, exposed to similar weather conditions, the system could relate these damage locations to the current bridge.

1.2 Need for IT Systems in the Field

Field workers, such as those described in Section 1.1, oftentimes have to perform their tasks in harsh environments. These tasks require access to various types of information, such as drawings and manuals, and oftentimes require that the field workers enter data about their processes. Some of the tasks described in Section 1.1 require the inspectors to access and interpret sensor information, to control inspection devices and to collaboratively solve problems with other coworkers distributed about a site. Thus, many of these field personnel are knowledge workers, in addition to performing physical labor-intensive activities. The inability to access and enter information during a field operation influences the efficiency and quality of that operation. Thus, while accessing and collecting information and applying their knowledge, these field personnel also need to have their hands free in order to be able to safely and efficiently move about the site on which they are working.

These field personnel need support for entering and accessing information as they perform their tasks. They cannot be standing there searching through pages of information displayed on their computer; it is both unsafe and impractical. To provide a safe and usable level of computing support in the field requires: 1) a hardware system that is unobtrusive and designed based on the field workers' tasks and mode of operation; and 2) a software system designed to support their specific tasks, provide knowledgeable advice about these tasks and provide intelligent, efficient forms of human-computer interaction that takes account of the field context of inspection.

The electronic technology available today allows the hardware to be smaller in size, while still possessing nearly the same capability as desktop or laptop computers. The wearable computer begins to provide users with the computing support they need in the field as they conduct other operations that require full use of their hands. Several generations of the wearable computer have been developed at Carnegie Mellon [CMU 99]. In most Carnegie Mellon applications of wearable computers, the emphasis has been on mobility, mostly hands free operation and either recording or displaying technical information. Commercial wearable computer products are now available. The existence and the availability of these wearable computers begins to address the need for unobtrusive hardware, though smaller, more energy efficient devices are still needed. However, the development of the software is still a difficult, time consuming task.

1.3 Organization of Paper

This paper first briefly describes a couple of recently developed field support systems. It then presents the various issues surrounding the development of IT support systems for field personnel. Finally, this paper describes an ongoing effort to create a framework that enables the developers of speech-controlled field data collection applications to develop their applications faster and more efficiently.

2.0 SEVERAL RECENT IT SYSTEMS FOR FIELD PERSONNEL

To date, most speech-based field data collection applications that have been tried and used in the field have supported command and control and data entry capabilities. Several industries and government agencies have begun to employ wearable computers to help their inspectors, including the airline industry, automobile industry, etc. Some of these applications provide speech interfaces for their users. Two field data collection systems have recently been developed into which a speech interface has been integrated. The first application is a system that supports bridge inspectors in the field [Garrett 98] and the second is a system for collecting construction progress monitoring data [Reinhardt 99].

MIA (Mobile Inspection Assistant), shown in Figure 1, is a wearable computer system that helps bridge inspectors to collect multimedia information in the field and produce their inspection reports. MIA allows an inspector to fill out an inspection form, access previous inspection reports, make sketches of the bridge elements, take photographs, and compile inspection reports via a speech or pen interface. We have successfully integrated a speech recognition capability into the system allowing the inspector to navigate through the list of bridge elements and entering the inspection findings, such as defect type, extent, and location, using speech input.

Progress Monitoring is a wearable computer system that allows construction site inspectors to collect information about the progress of the construction projects they are managing [Reinhardt 99]. The software enables the user to query elements that are related to late activities, current activities, activities under way, critical activities, and activities that have not been updated for a certain period of time. The software also allows the user to make comments that relate to the activities and elements associated with a construction project. The system supports both a speech and a pen interface.

3.0 ISSUES IN DELIVERING IT SYSTEMS IN THE FIELD

There are several major steps in the development of an IT system for the field. First, one must understand the field worker's context. Then, the appropriate functionality for that context needs to be designed. Next, appropriate hardware, interfaces and software components need to be integrated into a system. Finally, the system has to be extensively field tested in the actual field context with the actual end users, the field workers. The issues associated with each of these steps are described in the next six subsections.

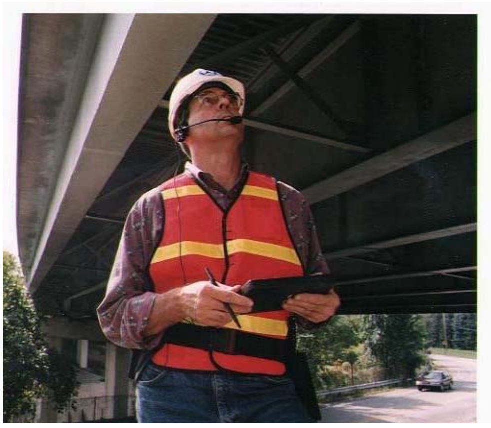
3.1 Understanding of Field Worker's Context

The first step in developing an IT system for field workers is to understand the context in which the field work is being performed. One needs to know:

- the type of work being performed in the field, such as condition assessment, which involves some preparation of the element to be assessed, such as measurements and calculations;
- the nature of the work environment, such as whether the work is performed year-round in all weather or only in warm, dry weather, or whether the site is extremely noisy;

- the type of information needed in the field and how they use that information; and
- the spatial and mobility constraints, such as whether the field worker will have to be climbing constantly into small tight spaces, or whether the person would be able to occasionally stop and concentrate on the IT device.

This type of information can only come from visiting the site, and observing and interviewing the field workers. In addition, a scenario describing the workers context should be developed and presented to potential users of the IT device to get their feedback on the developers understanding of the their work context.



b) MIA Hardware: 180 MHz ViA System with noise canceling microphone

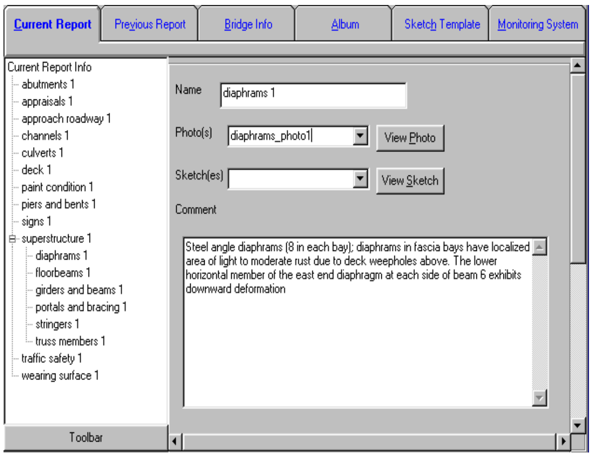


Figure 1. MIA Hardware and User Interface

3.2 Functionality

The next step in the development of these devices is to determine what specifically the field workers want to do, or could do, with these devices. Developers must determine the IT support needed, such as data entry support, decision support, and communications. Also, they must determine the type of human-computer interactions that can be effectively used by the field personnel in this context, such as button-based menu selections, pen-based hand-writing recognition, or speech recognition.

To properly support a field worker, the hardware and software must be designed to provide intelligent, effective and usable functionality for supporting the tasks being performed. For example, for bridge inspectors, the necessary information about a bridge, at the right level of detail, must be collected and organized ahead of time, as it is now in their paper-based approach, and entered into, or made accessible from, the computer. Some of this information will indicate where damage has been previously located, while other information may be recommendations about where to look on the bridge for damage, such as fatigue-prone details. Intelligent support is needed to assist the inspector in finding the needed information quickly and effectively while in the field. It is not enough just to put a desktop-oriented web site, supported by a one of the current search engines, in the hands of field personnel using a laptop or a palmtop computer. Field inspectors cannot be standing there brute-force searching through pages of possible “hits” displayed on their computer. Such activity is both unsafe and impractical.

In all field applications, if the field worker needs to put too much effort into using the device, which might be the case if the functionality provided by the system is complicated and more suited to a desktop environment, then the quality of the tasks they perform in the field will be seriously impaired. Thus, the system developer has to be extremely careful to design functionality that is truly needed and useable, in the field context. Adding additional unnecessary functionality may then force undesirable user interactions for the field context, such as scrolling and selections from long menus.

3.3 Mobile Computing Hardware

Commercially available, mobile, wearable CPUs with significant computing power are emerging from companies, such as Xybernaut, ViA, IBM, and Compaq. For example, the MA IV device from Xybernaut delivers a fully functional Windows 98 environment, based on an Intel Pentium chip with a clock speed of 233MHz. For one of the early prototypes of MIA, shown in Figures 1a and 1b, we used a ViA wearable computer based on a 180 MHz Cyrix chip. Much thinner clients are also available from companies like Palm, Handspring, and others, for supporting much more specific, predominately text-based functionality. If speech recognition, or some other computationally intensive functionality, is desired, larger amounts of computing are needed to support the speech processing. One means to provide this computing power is to wear it locally. For example, the Xybernaut MA IV provides sufficient power for high quality, mobile speech recognition. As wireless communication technology becomes more sophisticated, other options for delivering sophisticated functionality to a field worker are possible. For example, a smaller, less expensive device can communicate wirelessly with a more powerful base computer providing the sophisticated computing services needed.

3.4 Interface

One of the key issues in delivering an IT system in the field is providing a useful and usable interface between the field worker and the computing support. A variety of different types of visual displays are now available. The Xybernaut MA IV can be purchased with either a head-worn VGA display or a portable, flat panel VGA display that can be either hand-held or arm-worn (see Figure 2a and 2b). For the first MIA prototype, we attempted to develop a chest-worn display (see Figure 2c) that would allow the inspector to interact with the display in a nearly hands-free manner (through speech recognition) and then fold up the display when not in use.

While in the field, traditional input devices such as a keyboard and touch screen, become inappropriate to use because of their size and the way they were designed to be used. Pen-based systems, while appropriate for some field tasks, still do not address the need for hands-free operation in most inspection contexts. Other interaction methods then have to be explored to determine if they can be effectively used in a field environment. Again, the key here is that the form of human-computer interaction has to be unobtrusive, allowing field workers to move freely from point to point. The inspector must not have to focus heavily on operating the device. If the field inspector needs to put too much effort into operating the device, the quality of the tasks they perform will be impaired due to either stress or fatigue [Nasbaum 95].

a) Head-worn display b) Hand-held display c) Chest-worn display

Figure 2. Different Mobile Visual Displays

One approach to providing an efficient form of human-computer interaction is through a speech interface, where the majority of interaction is through speech and audio feedback, though interaction with Graphical User Interfaces (GUI) is also needed. However, developing and employing such an interface is time consuming, and requires a great deal of effort due to lack of understanding in developing and employing such interfaces.

The basic concepts of the speech recognition can be found in the literature [Cohen 94, Markowitz 97, Nasbaum 95, Rodman 99, Simpson 85]. The characteristics of a speech recognition system can be categorized using three dimensions, namely flow of speech, speaker dependency, and vocabulary size. First, the system can either be a continuous system, which allows the user to speak to the system continuously without pausing between each word. For a discrete system, the user has to stop between each word while dictating. Secondly, the system can be speaker dependent, where the system requires the user to provide the system a sample of speech in order for the system to adjust its parameters to recognize the user's speech (training). For a speaker independent system, the user is not required to train the system. Finally, with respect to the size of the vocabulary, some systems can recognize up to 60,000 active words, while some systems can only recognize the 10 individual digits.

The usage of speech recognition in an application can be divided into four categories: dictation, command and control, data entry, and information retrieval. In dictation, the user of a speech recognition system simply speaks into a microphone and the speech recognition system translates the signal into words that are sent into a program, such as a word processor, to create a document. Command and control deals with using speech to control a particular device or a software program, e.g. using speech to navigate through different parts of a program. Data entry is more or less the same as a dictation application except that the vocabulary is somewhat more limited. Finally, information retrieval systems allow users to ask for a specific piece of information using natural spoken language.

The task of the speech recognition engine is to transform the speech input into text. The speech recognition matches the acoustic signal from the speech to those of words in the vocabularies; therefore, only words in the vocabulary are capable of being recognized.

As one can see, vocabularies play an important role in speech recognition, since the vocabularies define the set of words or phrases that can be recognized by a speech recognizer. Speech grammars are an extension of the single words or simple phrases supported by vocabularies. They are structured collection of words and phrases bound together by rules that define the set of speech streams that can be recognized by the speech recognizer at a given point in time. This is important because the smaller the vocabularies that need to be recognized at a given point in time, the more accurate the results will be.

Thus, in developing a speech-based application, one has to carefully select the type of speech interaction, specify the vocabulary, and design the grammars that will be used for that application.

In addition to the computing hardware and the user interfaces, additional issues arise related to providing sufficient power for the device. Field workers are not usually co-located with a convenient source of electric power. The battery technology is getting better, but most commercial platforms still require the field worker to carry extra batteries. In most cases, the device has to be shut down when the batteries are to be swapped. The Xybernaut MA IV does allow for "hot-swapping" of the batteries, but the device must be put into a mode that is not much more convenient than a complete shutdown. For the foreseeable future, providing sufficient and convenient power sources for these mobile devices in the field will be a subject of research and development. However, this issue will be greatly helped by the battery research being done in the mobile phone industry. For example, Motorola is researching the possibility of using refillable fuel cells for long-term power for their cell phones and computers. [Reuters 00]

Advances in wireless communications have made it possible to easily link field workers with each other and centralized compute and data servers. We have experimented with the use of Lucent's WaveLAN spread-spectrum PCMCIA wireless LAN card in both the Xybernaut MA IV and the ViA to communicate with other computers either having another WaveLAN cards or a Lucent access point.

3.5 Software

To date, our software has been fast prototyped using Microsoft Visual Basic as the means to develop the interface views and controls presented to users in the field. Where possible, we have attempted to practice object-oriented development strategies. We have also attempted to use existing commercially available software components, such as available speech recognition engines, wherever possible. One of the issues we identified early on in the development of software for IT devices to be used by field workers is that the software functionality must be customized for the specific task being supported. The software must take account of the field worker's tasks being performed in the field and the types of interactions they are prepared to have with the field IT device. In other words, it is not possible to have a "one size fits all" software architecture and set of components that will meet the IT needs of field workers. One issue we have identified for future research is the need for advanced software architectures and development frameworks to support the development of these IT systems for field workers.

3.6 Field Evaluation of System with Users

The effect of the physical, field environment plays an important role on how a human operator interfaces with the mobile IT device in the field (e.g., speech utterances produced in the field differ from those made in a lab setting). Another important factor that affects the performance of field workers using the IT device is the application environment, which consists of the tasks and application for which the device is being used. For example, when delivering a speech recognition interface for a device to be used in the field, it is very important to choose tasks and applications that are appropriate for speech recognition [Baber 91, Nasbaum 95]. For example, it is well known that speech is not a good interface for drawing an object. Therefore, in bridge inspection, one likely will not want the bridge inspectors to use speech to create or edit a sketch.

As illustrated above, to create IT devices that can be effectively and efficiently used in the field requires an understanding of the technology as well as the context of the problem domain. Unfortunately, there are no general rules of thumb that can be employed when one wants to apply different hardware and interface technologies, such as speech recognition technology, to a particular application. Since general criteria for the design of these devices are not yet completely understood, a true understanding of the advantages and limitations of the device and its interfaces now can only be obtained by prototyping and field-testing it with actual end users. Many ideas and interactions that were felt by the developers, and some end users involved in the early design stages of the device, to be potentially useful and effective interactions turn out to be problematic in the field. For example, the chest-worn display was felt at first to be "the best solution" for delivering access to the display while keeping the bridge inspectors hands free. What we discovered in the field was that the inspectors found the display to be hard to use, especially when they did want to use their hands to mark on a drawing on the screen. The inspectors could not easily maneuver on the screen when it was so close to their chest.

4.0 DISCUSSION: WHAT ASSISTANCE IS NEEDED?

Based on experiences with designing, building, and field testing mobile computing systems for use by field workers, there is a great need for development frameworks to assist developers in rapidly prototyping and field-testing hardware configuration and interface designs.

For example, both applications described in Section 2.0 integrate two speech recognition features: a command and control use of speech recognition that allows the inspector to navigate through different parts of the system, and a data entry feature that allows an inspector to input their findings. These two forms of speech support are the most basic. We found during the field testing of the MIA system that the users had to focus more heavily on how the system behaves than on the tasks they were doing. The inspectors had to closely monitor the feedback from the system due to speech recognition errors. The vocabulary and grammar had to be improved. However, errors from the speech recognition system could not be completely eliminated. Therefore, we discovered through prototype modification and field testing that the system also had had to provide a highly efficient error correction mechanism.

To better utilize a speech interface, different forms of interaction mechanisms need to be explored. For example, it might be better to provide a dialog management system for a bridge inspector during inspection. The dialog has to be able to handle errors committed by the user and issue prompts asking for appropriate clarification from the user. However, integrating a speech dialog is not a simple task. A framework for helping the application developer in incorporating different speech interaction mechanisms is obviously needed. With such a framework, developers will be able to more quickly generate prototypes for field testing interaction mechanisms.

Our research group at Carnegie Mellon is developing frameworks for assisting the developers of audio-centric data collection applications. One of the frameworks being developed is composed of two major components: a Speech Manager (SM) and a Data Collection Application Manager (DCAM). The DCAM will define the control function for a data collection application. It will handle all common activities in the data collection process, such as presenting the previous information about a specific inspected object, activating and deactivating the appropriate data collection task, updating new inspection data gathered during the inspection process, etc. The DCAM will be developed from scratch taking advantage of our experience with developing field data collection applications. The SM will encapsulate all the functions that an application developer needs to be able to integrate a speech interface, including a speech recognition engine, grammars, speech prompts, dialog manager, etc. By dividing the framework into two components (SM and DCAM) that separate the speech functionality from the data collection functions, we will be able to make use of the existing object-oriented speech regognition frameworks (such as SpeechObject [Vcom 99]) by modifying them to support our requirements. By keeping the SM separated from the DCAM, application developers will be provided a much more flexible development environment. They will be able to build their application on top of the SM only, the DCAM only, or both. Finally, by maintaining separation, other framework components, such as a framework for developing gesture-based interaction can be more easily integrated into the overall framework.

5.0 CLOSURE

Field workers need support for entering and accessing information as they perform their tasks. Since general criteria for the design of these devices are not completely understood, a true understanding of the advantages and limitations of the device and its interfaces currently can only be obtained by prototyping and field-testing it with actual end users. Based on experiences with designing, building, and field-testing mobile computing systems for use by field workers, there is a great need for development frameworks to assist developers in rapidly prototyping and field-testing hardware and user interfaces. Our research group at Carnegie Mellon is currently developing frameworks for assisting the developers of audio-centric data collection applications for supporting field workers in civil engineering-oriented contexts.

6.0 ACKNOWLEDGEMENTS

We thank PITA (Pennsylvania Infrastructure Technology Alliance) for providing funding for the development of the Mobile Inspection Assistant (MIA). We thank the Michael Baker Corporation for their assistance and feedback in the development of MIA. We thank the bridge inspectors of PENNDOT District 11 for their assistance in field testing MIA. We also acknowledge the work by Mr. Jan Reinhardt, who developed the Project Monitoring prototype briefly mentioned in this paper and described more fully in a separate paper presented at this conference.

7.0 REFERENCES

[Baber 91] C. Baber. "Task Selection for Automatic Speech Recognition in Control Room Systems." *Speech Technology*, Feb./March 1991, pp. 82-85

[CMU 99] "Wearable Computer Systems at Carnegie Mellon University." <http://www.cs.cmu.edu/People/wearable/frontpage.html> 1999.

[Cohen 94] P. R. Cohen, S. L. Oviatt. "The Role of Voice in Human-Machine Communication." *Voice Communication Between Human and Machines,* D.B. Roe and J. G. Wilpon (Eds), pp. 34-75, 1994

[Garrett 98] J. H. Garrett, Jr. and A. Smailagic. "Wearable Computers for Field Inspectors: Delivering Data and Knowledge-based Support in the Field, *Proceedings of the Workshop on AI in Structural Engineering*, ed. Ian Smith, pp. 146-164, Centro Stefano Franscini, Monte Verita, Anscona, Switzerland, July 1998

[Johnson 88] R. Johnson, B. Foote. "Designing Reusable Classes." *Journal of Object-Oriented Programming,* Vol 1, No 2, pp. 22-35, June/July 1988

[Markowitz 97] J. A. Markowitz. *Using Speech Recognition*, Prentice Hall, 1997

[Nasbaum 95] H.C. Nasbaum, J. Degroot, L. Lee. "Using speech Recognition Systems: Issues in Cognitive Engineering." In *Applied Speech Technology*, CRC Press, 1995

[Nuance 99] "Nuance SpeechObjects and V-Commerce Applications." *A Nuance white paper*, Nuance Communications, 1999

[Reinhardt 99] J. Reinhardt. "Progress Monitoring with wearable computers." 1999 <http://www.ce.cmu.edu/~janr/icmms/icmms.html>

[Reuters 00] Reuters News Service. "Motorola Makes Fuel Cell to Power Laptops, Phones." January 19, 2000. <http://216.33.41.203/vtech/011900/fuelcell.sml>

[Rodman 99] R. D. Rodman, *Computer Speech Technology*, Artech House, 1999

[Simpson 85] C. A. Simpson. "System Design for Speech Recognition and Generation." *Human Factors*, 27, pp. 115-141, 1985

[Srinivasan 98] S. Srinivasan, J. Vergo. "Object Oriented Reuse: Experience in Developing a Framework for Speech Recognition Application." *In Proceedings of the 1998 International Conference on Software Engineering*, pp. 322-330, IEEE, 1998

[Vcom 99] "V-Framework." <http://www.v-commerce.com/>