

A MULTIMODAL CONTEXT AWARE MOBILE MAINTENANCE TERMINAL FOR NOISY ENVIRONMENTS

Fredrik Vraalsen, Trym Holter, Ingrid Storruste Svagård, and Øyvind Kvennås

SINTEF ICT, N-7465 Trondheim, Norway

Abstract: Maintenance workers in the oil and process industry have typically had minimal IT support, relying on paper-based solutions both for the information they need to bring into the field and for data capture. This paper proposes *a mobile context aware system for maintenance work* based on electronically tagged equipment and handheld wireless terminals with a multimodal user interface. Particular attention has been given to voice interaction in noisy industrial scenarios, utilising the PARAT earplug. A proof-of-concept demonstrator of the system has been developed. The paper presents the demonstrator architecture and experiences gained through this work.

Key words: Mobile systems; maintenance; multimodality; context awareness; noise robustness; automatic speech recognition; electronic tags.

1. INTRODUCTION

The process industry has implemented elaborate maintenance procedures to maximize production throughput and minimize downtime due to uncontrolled failures. Estimates show that for critical equipment, unplanned corrective maintenance costs on average three times more than planned maintenance work. Furthermore, estimates show that in process industry today, only about 65% of the total maintenance work hours are spent on actual repair work. The remainder is spent on preparations, searching for documentation, applying and waiting for work-permits, reporting etc. (SINTEF 2002). Another problem reported by our user group, described in chapter 2, is that the documentation of historical maintenance work is often

not satisfactory. This is attributed to the process itself being cumbersome, and that the workers lack the incentive for it. The potential for quality improvement and cost reductions related to maintenance work seems obvious.

Mobile and wearable computers supporting maintenance and inspection tasks have been subject to attention from both researchers and suppliers for some years, such as for aircraft inspections and maintenance (Ockerman and Pritchett 1998)(Siegel and Bauer 1997), train maintenance (Siewiorek et. al. 1998) and maintenance of buildings and technical installations (Ailisto et. al. 2003)(Stäger et. al. 2003). A common focus has been increased efficiency and accuracy compared to paper-based approaches. Other focus areas include hands-free operation through the use of voice commands and heads-up displays (Ailisto et. al. 2003), context aware systems (Ockerman and Pritchett 1998), self-diagnostic systems and sensor networks (Stäger et. al. 2003) and remote collaboration through the use of wireless communication (Ockerman and Pritchett 1998)(Siewiorek et. al. 1998).

This paper presents a scenario based on the process industry. We propose a mobile maintenance support system designed specifically with a focus on multimodal user interfaces (Comerford et. al. 2001)(Huang et. al. 2001) suitable for hands- and eyes-free operation. Such interfaces offer the user freedom to use a combination of modalities depending on the specifics of the task or environment (Oviatt et. al. 2000). An unsolved problem in this field is the fragile performance of speech recognition in noisy environments (Cole et. al. 1995), making deployment in process industry challenging due to the potentially severe noise levels (Tempest 1985). This problem is addressed through the use of the PARAT earplug (Aakervik 2000)(NACRE 2004), an intelligent device designed for hearing protection, speech recognition and human-to-human communication in noisy environments.

The approach proposed here also utilises context information from electronic tags on equipment to automate equipment identification and facilitate onsite access to data in a backoffice maintenance system through wireless communication.

The remainder of this paper is organised as follows: Chapter 2 discusses the user requirements that form the background for the system design. A scenario and an overview of the proposed mobile system are presented in chapter 3. Chapter 4 describes the design and implementation of the proof-of-concept demonstrator. The paper closes with a discussion in chapter 5 and conclusions and future work in chapter 6.

2. USER REQUIREMENTS

The maintenance scenario and related user requirements were studied through a series of informal meetings with offshore personnel from a major oil company that were experienced with maintenance, both managers and field workers. The following were perceived as the most important areas for improvement:

1. Access to data: Retrieval of background maintenance data often involves searches in large paper-based archives. This process is time consuming, and too often the required material can not be found.
2. Quality of maintenance data: The quality of historical maintenance data is often substandard, and such data is thus of little help. One manager stated that "solutions that could lower the documentation threshold *just a little bit*, and make the workers report *some* more of their findings, would be a big improvement".
3. Information flow and work-processes: Work-orders and work-permits are most often paper-based. This process often introduces long delays as the documents travel between the involved parties.
4. Access to personnel: When a situation requires expert support, delays occur while waiting for such help. If advice could be given from a remote location, these delays could be avoided. In an offshore setting, this even opens for locating key personnel onshore, from where they could support multiple installations.

Based on this input, use cases were developed describing detailed user-system interaction and corresponding system response for a range of activities. The use cases were presented to the users for feedback and the system design was adjusted accordingly. The most important aspects of the user feedback can be summarized as follows:

- Mobile access to maintenance data such as work and safety procedures is a good idea. It is particularly important as it is a trend, especially offshore, that the personnel must be able to service an increasing number of different types of equipment, and the maintenance worker may not necessarily be an expert on all the equipment that he or she services.
- Many maintenance workers have little experience with mobile terminals. Thus particular emphasis must be put on usability and reliability aspects of the solutions. Including well known functionality and user interfaces (like mobile telephony) into the device could be a way to increase user acceptance.
- Hands-free operation of the device is required to make it simpler to document and report findings for work situations where the hands and/or eyes are busy.

- Electronic tagging of equipment is perceived as a good idea for several reasons. Most importantly it simplifies the equipment identification process and is less error prone and cumbersome than manual identification.
- The maintenance application must allow the worker to document irregular findings whenever and wherever he or she discovers them.

There are a number of user requirements beyond this, which our user group did not focus on. Most important among these were information security and safety requirements, such as ATEX (ATEX 1994). In our proof-of-concept demonstrator we chose not to focus on these issues. Human-to-human communication aspects (e.g., remote collaboration with experts) were also kept outside of the scope of work.

3. SYSTEM CONCEPTS AND OVERVIEW

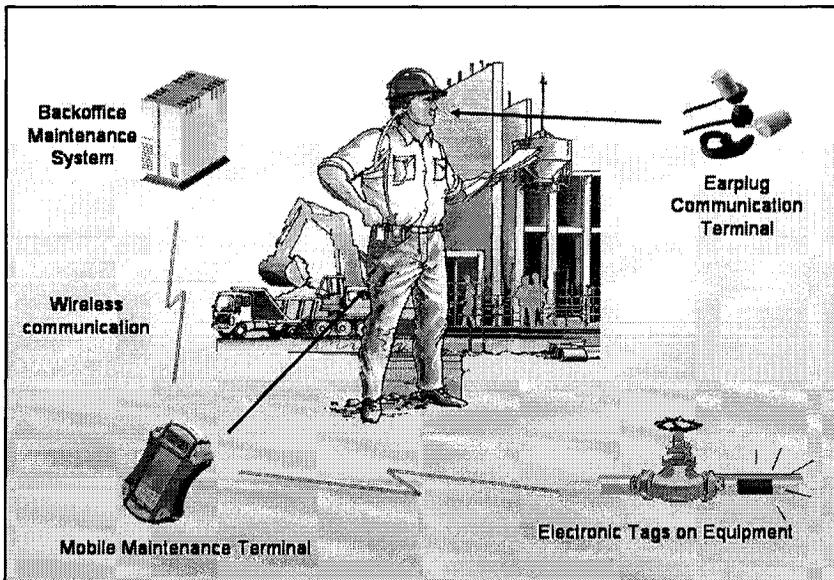


Figure 1. Conceptual overview

Figure 1 shows a conceptual overview of our suggested maintenance system. The following is a scenario that illustrates the system vision:

Maintenance worker Baker approaches gearbox GB-13-407 in area A30 for a routine inspection, in accordance with the work order just downloaded to his mobile maintenance terminal. As he approaches, his terminal communicates with the radio tag on the gearbox and verifies that Baker is at

the right device. Before Baker can start the inspection, the system prompts him to read through a new safety procedure that recently has been implemented. He also takes a brief look at the maintenance log, downloaded in real-time from the backoffice maintenance system. Before he rolls up his sleeves and goes to work, he activates the application's speech mode by tapping the screen and puts the terminal in his belt. Audio prompts received through the earplug terminal guides him through the inspection procedure, one step at a time. The earplug eliminates the strong background noise and captures his commands and annotations flawlessly, allowing him to navigate through the check list and document his work with his hands free. At one point during the inspection he requests to see a technical diagram for guidance. He temporarily halts the procedure and inspects the diagram on the screen, before returning the terminal to his belt. On completion of the work, a report is compiled. Baker signs the report and is ready for the next work order on his list.

The rugged mobile terminal runs a dedicated maintenance application which communicates wirelessly with the backoffice systems. Electronic tags labelling the equipment provide context information and enable the application to present the user with information and services relevant to the current maintenance task (Dey 2001). The user can interact with the application in different ways based on preference or requirements of the current environment or task, and noise-robust speech recognition solutions facilitate hands-free operation even in harsh industrial environments. Services offered by such a system would typically include:

- Administrative tasks, such as work scheduling
- Display information about equipment, e.g. design schematics and maintenance logs
- Work support, interactively guiding the user through the maintenance process
- Log the performed tasks and measurements
- Automatically generate and transmit maintenance reports and notifications
- Cooperative maintenance work, e.g. on-line communication with experts

The system can also help improve safety by ensuring that procedures are followed, e.g. by handling work permits and including safety steps in the interactive maintenance guide.

The UML class diagram in Figure 2 shows a platform independent¹ view of the main maintenance system components. The maintenance application consists of components performing tasks such as generating reports and identifying equipment based on the tags. It can also contain a number of user

¹ Independent of implementation strategy, e.g. programming language or physical distribution

interface components, which may in turn rely on other components, e.g. for speech recognition and text-to-speech synthesis. A controller orchestrates the various components.

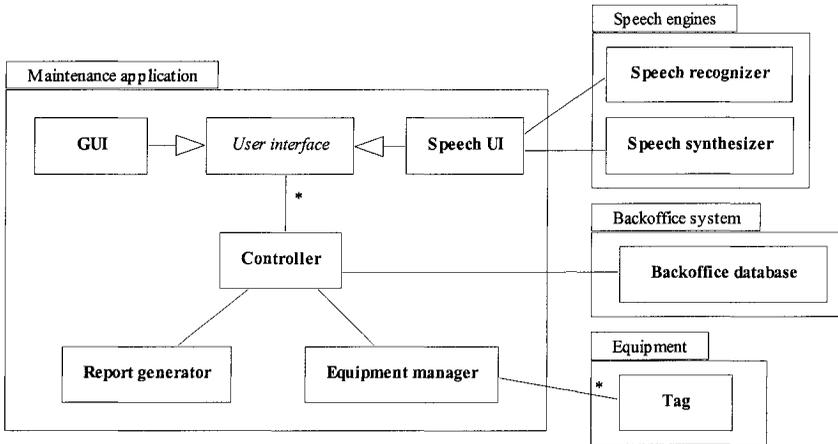


Figure 2. Main system components

4. SYSTEM DESIGN AND IMPLEMENTATION

A demonstrator was implemented to show the core functionality of the maintenance system described in chapter 3. The main components of the demonstrator, their interrelationships, and their physical distribution are shown in the UML deployment diagram in Figure 3. The diagram is a platform specific realisation of the design in Figure 2. The remainder of this chapter discusses aspects related to the implementation of this design.

4.1 Development environment

The mobile device we have chosen for the demonstrator is an iPAQ 5450 PDA, as it contains most of the desired features, most notably built-in WLAN (802.11b) and Bluetooth for wireless communication and equipment identification, as well as support for connecting an external microphone. In addition, the iPAQ provides ample expansion capabilities through the use of various expansion jackets. The iPAQ is running Pocket PC 2002.

The mobile maintenance application has been implemented mainly in Java, using the Java 2 Micro Edition (J2ME) Personal Profile, and is running on the IBM J9 Virtual Machine. Some low-level features are implemented in

C/C++, described in more detail below. The backoffice maintenance system is a simple mock-up to allow testing of the mobile maintenance application. It consists of an Apache web server which provides the maintenance application with maintenance data, e.g. work orders and equipment technical data, and a simple server that receives and processes maintenance reports.

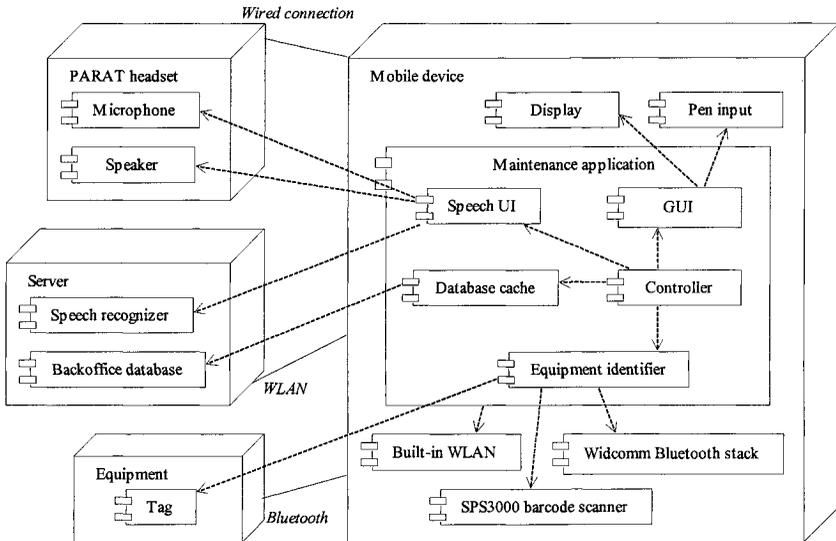


Figure 3. Maintenance system components and their physical distribution

4.2 Demonstrator functionality

The maintenance application initially displays an overview of current work orders. Selecting a work order, the user can view the work order details, such as maintenance procedure descriptions and equipment technical data, as shown in Figure 4, in addition to performing administrative tasks such as checking and requesting work permits. Once the user has selected a work order, he or she can start the inspection. The interactive maintenance guide is similar to a wizard dialog, consisting mainly of a sequence of questions with predefined answers, illustrated in Figure 5. The user can also attach comments to each maintenance inspection step in the form of free-text comments or audio recordings. Once the inspection is completed, the user is presented with a report summary for confirmation before sending the report to the backoffice maintenance system via the wireless network.

The maintenance application offers a multimodal user interface with support for speech-based input and output in addition to a GUI with pen-

based input. The user can switch freely between the two modalities, allowing him or her to select the most suitable interaction method. One of the advantages of using speech input is that the user has both hands available for other tasks. This is even more important due to the small size of mobile devices, which often means that tactile input is cumbersome to use. The device also offers speech output for certain parts of the application. It is sometimes argued that speech output is impractical due to the amount of information (Ailisto 2003)). We find however, that during the wizard dialog, the amount of information is fairly small, and thus speech output is advantageous because it also allows *eyes-free* operation of the application.

One of the main obstacles to using interfaces based on speech technology is the fragile performance of speech recognition in noisy environments (Cole et. al. 1995). In the next section we describe how we aim to mitigate this problem.

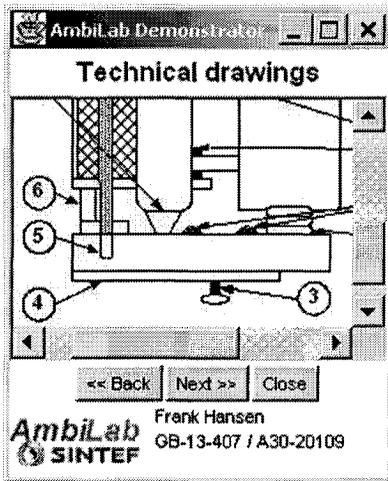


Figure 4. Equipment technical data



Figure 5. Inspection wizard

4.3 The PARAT communication terminal and automatic speech recognition

PARAT (Personal Audio/Radio Active Terminal) (Aakervik 2000)(NACRE 2004) is an intelligent hearing protector and audio interface for communication systems. A principle sketch is shown in Figure 6. An earplug with a seal sits in the ear. In a quiet environment the sound captured

by the outer microphone is reproduced unaltered by the loudspeaker on the inside of the earplug. When the noise becomes severe, the transmission from the outer microphone is processed and gradually reduced.

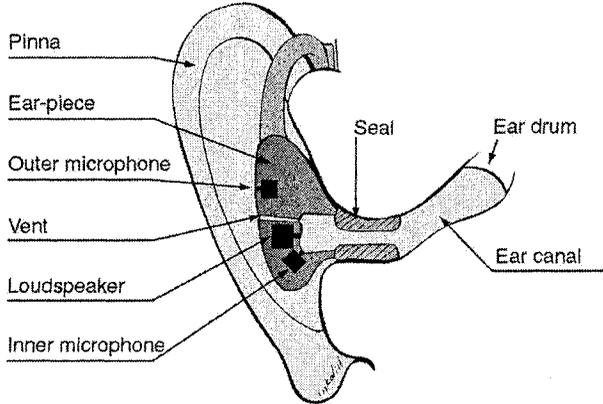


Figure 6. The PARAT earplug

The primary purpose of the inner microphone is to capture the user's voice in a manner that is robust against even strong ambient noise, and can be used both for voice communication and speech recognition purposes. This approach to speech recognition is complementary to the main-stream research in noise-robust speech recognition over the last few years, and has been shown to be successful (Strand et. al. 2003). Noise-robust speech recognition research has to a large degree been focusing on signal and feature compensation methods, driven through the ETSI Aurora initiative (ETSI 2004). The approach described here on the other hand focuses on the signal acquisition level.

For speech recognition we use Sprex ANSR (Sprex 2004), a network-distributed speech recognition system. Some initial processing of the speech input is done locally, e.g. to identify conversation end-points, before the processed data is transferred over the wireless network to a server where the actual recognition is performed. The results, i.e. the recognized phrases, are then returned to the client where they are used to generate the relevant UI events.

Given the fairly simple dialog structure and limited number of speech prompts, we have chosen to use pre-recorded speech for speech output. This gives better speech quality than a text-to-speech synthesis running on the PDA or on the network, of course at the cost of lower flexibility. The pre-recorded speech files are stored locally on the PDA.

4.4 Tag solution

In our application, tags serve the purpose of identification only; hence no data will be stored in the tags except for the ID of the object. Thus several tag types could be applied; both passive (no battery) and active (battery-powered) tags would serve the purpose.

The actual implementation offers three different methods, illustrating the choice and flexibility possible among tagging schemes. The primary method is using the built-in Bluetooth communication to read the unique Bluetooth address of a Bluetooth tag, deriving the equipment ID through a database lookup. This method illustrates active tags and the fact that objects can be identified without visibility to the physical tag. This approach enabled us to use readily available technology (Bluetooth is embedded in the iPAQ 5450, Bluetooth tags are easily acquired and protocol stacks exist), hence reducing implementation time and risk. The drawback is the limitations of battery life time, requiring batteries to be replaced at regular intervals.

To overcome the potential problem of battery life time, the demonstrator also supports scanning barcode-based IDs using the Symbol SPS3000 barcode scanner. Bar code scanning is an extremely cheap and extensively used technology, and virtually no maintenance is required. However, bar code scanning requires visibility which may be hard to satisfy in an industrial process plant.

Finally, as a supplement to Bluetooth and bar code scanning, the user can manually enter the ID. This is important as a fallback solution, both because deployment of new technology may be gradual and electronic tags may thus not always be available, and for robustness in case the Bluetooth or barcode solutions are not working.

Although not implemented in the demonstrator, other tagging methods would also be adequate in our application, in particular RFID, which solves both the problem of battery life time and visibility. RFID tags will be the subject of future extensions of the demonstrator.

4.5 Communication with backoffice system

The maintenance application uses the built-in WLAN to communicate with the backoffice maintenance system. As a constant network connection is assumed by the chosen speech recognition solution, there is no functionality that enables off-line work for the user in the current demonstrator. However, we do cache information related to the current work task locally, e.g. technical diagrams, mainly for efficiency reasons. This is most likely not a reasonable choice in a real production setting, but in this

demonstrator the primary concern was to exhibit speech recognition and control.

XML (Extensible Markup Language) is deployed for information exchange between the backoffice system and mobile terminal. XML defines the structure of documents through the use of DTDs (Document Type Definitions) or XML Schema, but leaves the meaning or *semantics* up to the user and application. Practically all application content, i.e., work orders, inspection lists, equipment data, maintenance reports, etc., is defined in the form of XML documents with accompanying DTDs. This simplifies the procedure of adjusting the application in response to requirements changes and was found very useful. Furthermore, XSL (eXtensible Stylesheet Language) transformations can be used to translate between different XML formats and to other textual or binary data formats, facilitating integration with a wide variety of third party systems.

4.6 Implementation experiences

A number of low-level features of the PDA OS and hardware lack Java APIs in the J2ME Personal Profile, e.g. the Bluetooth stack. The AveLink Bluetooth Suite (*Atinav* 2003) provides a Java Bluetooth API (JSR-82 2004) implementation for J2ME, however this was not yet available for the iPAQ 5450 at the time of implementation. Instead, we used the Widcomm Bluetooth SDK (*WIDCOMM* 2003) for the iPAQ 5450 to implement a subset of the Java Bluetooth API using C++ and Java Native Interface (JNI). This solution provides us with a level of portability, allowing us to easily move the application to different platforms which provide the Java Bluetooth API. Similar solutions have been used for other low-level functionality, e.g. recording and playback of sound.

The use of Java and XML has made it simple to adapt the demonstrator to new and changing requirements. The portable nature of Java has also enabled a more rapid develop-test cycle, as we were able to develop and test most of the application (excluding the device specific functionality) on a regular PC before deploying it on the actual PDA. Testing on the PC was much faster, both as transferring updates to the PDA incurred an overhead and as the lower processor and communication performance on the PDA leads to more time-consuming testing. Furthermore, by utilising software design patterns such as MVC (Krasner and Pope 1998) and Mediator (Gamma et. al. 1995) we achieve a degree of decoupling which enables us to replace individual components without changing the rest of the system.

Accessing low level features through native code turned out to be a feasible workaround until implementations of the necessary Java APIs are

available on the PDA. However, much of the device specific functionality was not covered by the standard Windows CE/Pocket PC API, and finding and acquiring the necessary development documentation and SDKs for the device specific features can be both time consuming and costly, if they are available at all.

5. DISCUSSION

Although limited, our user requirements studies indicate that usability is a prerequisite for successful introduction of mobile work support in maintenance and operations. Similar observations have been done by previous studies (Ockerman and Pritchett 1998)(Siegel and Bauer 1997)(Ailisto 2003). The users focus on hands- and eyes-free operation, and speech technology is thus a key component. We believe that the suggested multimodal interface with the PARAT earplug has the potential to meet these requirements. Preliminary tests show that it improves noise robustness compared to standard solutions and its unobtrusive form-factor ensures that it will not interfere with the physical activities involved in maintenance work.

Tags were introduced into the system concept to provide context awareness. The user's position relative to the equipment is verified, giving automated and error-free identification of equipment. With radio tags in particular, this can be achieved with no active intervention from the user. Other approaches to positioning are often promoted for safety reasons, to monitor the whereabouts of personnel in critical situations. This has not, however, been the focus of our study.

Another advantage of active radio tags is the possibility of local data storage. One way to exploit this possibility is to store the maintenance log on the tag, which will always follow the equipment. This would be an advantage especially for equipment that is moved around, or even sent away for service and repair. The main drawback of active tags is the fact that they require power supply, and effective schemes for this must be devised before this is a realistic option.

Bluetooth tags were chosen to demonstrate the concept of radio tags in the demonstrator. The same functionality could have been achieved with RFID, but Bluetooth tags were chosen due to the easy availability of devices and components.

WLAN was chosen for wireless communication much for the same reasons. It also seems to be a realistic future choice in our application scenario, due to the low cost of infrastructure and use, high bandwidth, and satisfactory range. WLAN security has also improved over the recent years.

For scenarios where geographical distance is an issue, GPRS would in many cases be a realistic option.

Probably the most important issues related to introducing this kind of new technology, are the required organizational and procedural changes. For the technology to have the desired effect, the implications on all levels of the organisation have to be considered carefully. It is very unlikely that this process will succeed unless the users are motivated for the change, and the key will be to involve the users in the further stages of implementation.

6. CONCLUSIONS AND FUTURE WORK

This paper describes user requirements, design and implementation for a mobile maintenance support system targeting process industry environments. The feedback from a focus group working on the user requirements showed clearly the importance of hands-free operation of the mobile device, and emphasis has thus been put on noise-robust speech recognition, facilitated by the PARAT ear-plug communication terminal.

The proof-of-concept demonstrator described in detail in this paper is based on context awareness through electronic tagging of equipment and a mobile maintenance terminal with wireless communication with a backoffice maintenance server. Initial tests of this approach validate its usability.

In the next stage of this work, our goal is to have a pilot implementation and evaluation, based on further development of the demonstrator, in a real scenario and with real users. Potential application areas are offshore and onshore process installations, but also other kinds of applications areas will be considered.

REFERENCES

- Ailisto, H., Haataja, V., Kyllönen, V. and Lindh, M. (2003) 'Wearable context aware terminal for maintenance personnel', In: *Proceedings of The European Symposium on Ambient Intelligence (EUSAI 2003)*, Nov. 3-4 2003, Elsevier Science, Eindhoven, The Netherlands
- Cole, R. et. al. (1995) 'The challenge of spoken language systems: research directions for the nineties', *IEEE Transactions on Speech and Audio Processing*, vol. 3, Jan. 1995, pp. 1 – 21
- Comerford, L., Frank, D., Gopalakrishnan, P., Gopinath, R. and Sedivy J., (2001) 'The IBM Personal Speech Assistant', In: *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2001)*, May 2001, Salt Lake City, Utah
- Dey, K. (2001) 'Understanding and Using Context', *Personal and Ubiquitous Computing*, vol. 5, pp. 4-7
- Gamma, E., Helm, R., Johnson, R. and Vlissides, J. (1995) *Design Patterns – Elements of Reusable Object-Oriented Software*, Addison-Wesley

- Huang X. et. al. (2001) 'MIPAD: A Multimodal Interaction Prototype', In: *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2001)*, May 2001, Salt Lake City, Utah
- Krasner, G. E. and Pope S. T. (1998) 'A cookbook for using the model-view controller user interface paradigm in Smalltalk-80', *Journal of Object-Oriented Programming*, vol. 1, August/September 1998, pp. 26-49
- Ockerman, J.J. and Pritchett, A.R. (1998) 'Preliminary Investigation of Wearable Computers for Task Guidance in Aircraft Inspection', In: *Proceedings of The Second International Symposium on Wearable Computers (ISWC '98)*, Oct. 19-20 1998, Pittsburgh, PA, pp. 33-40
- Oviatt, S. et. al. (2000) 'Designing the User Interface for Multimodal Speech and Pen-Based Gesture Applications: State-of-the-Art Systems and Future Research Directions', *Human-Computer Interaction*, vol. 15, pp. 263—322
- Siegel, J. and Bauer, M. (1997) 'A Field Usability Evaluation of a Wearable System', In: *Proceedings of The First International Symposium on Wearable Computers (ISWC '97)*, Oct. 13-14 1997, Cambridge, Mass, pp. 18-22
- Siewiorek, D., Smailagic, A., Bass, L., Siegel, J., Martin, R. and Bennington, B. (1998) 'Adtranz: A Mobile Computing System for Maintenance and Collaboration', In: *Proceedings of The Second International Symposium on Wearable Computers (ISWC '98)*, Oct. 19-20 1998, Pittsburgh, PA, pp. 25-32.
- SINTEF (2002), *Offshore Reliability Data Handbook (OREDA)*, 4th ed, Det Norske Veritas, Norway
- Strand, O. M., Holter, T., Egeberg, A. and Stensby, S. (2003) 'On the feasibility of ASR in extreme noise using the PARAT earplug communication terminal', In: *Proceedings of IEEE Workshop on Automatic Speech Recognition and Understanding (ASRU 2003)*, December 2003, IEEE, Virgin Islands, USA
- Stäger, M., Junker, H., von Waldkirch, M., Tröster, G. (2003) 'Using Wearables in Maintenance: A Modular Test Platform', In: *Proceedings of TCMC 2003: Workshop on Wearable Computing*, Mar. 11-12 2003, Graz, Austria
- Tempest, W. (1985) Noise in Industry, in: Tempest, W. (ed) *The Noise Handbook*, Academic Press, New York, pp. 179-194
- Aakervik, A.-L. (2000, Dec. – copyright date) "An earplug passes the word", (*Gemini*), Available: http://www.ntnu.no/gemini/2000-06e/09_1.htm (Accessed: 2004, May 25)
- "ANSR – Action-oriented, Network-distributed Speech Recognition", (2004 – copyright date), (*Sprex*), <http://cassandra.sprex.com/ansr/>, (Accessed 2004, May 25)
- "ATEX Directive 94/9/EC: Equipment and Protective systems intended for use in potentially Explosive Atmospheres", (1994, Mar. 23 – copyright date), (*European Union*), Available: <http://europa.eu.int/comm/enterprise/atex/> (Accessed 2004, May 25)
- "Aurora – Distributed Speech Recognition", (2004, May 25 – last updated), (*ETSI*), Available: <http://portal.etsi.org/stq/kta/DSR/dsr.asp>, (Accessed 2004, May 25)
- "aveLink Bluetooth", (2003 – copyright date), (*Atinav*), Available: <http://www.atinav.com/bluetooth/index.htm> (Accessed 2004, May 25)
- "JSR-82: Java APIs for Bluetooth", (2004 – copyright date), Available: <http://www.jcp.org/en/jsr/detail?id=82> (Accessed 2004, May 25)
- "Natural Communication in Rough Environments", (2004, Jan 30 – last updated), (*NACRE*), Available: <http://www.nacre.no> (Accessed 2004, May 25)
- "WIDCOMM: Bluetooth Connectivity Solutions", (2003 – copyright date), (*WIDCOMM*), Available: http://www.widcomm.com/Products/bluetooth_comm_software_btce.asp (Accessed 2004, May 25)