UCD in Wearable Computing – A Case Study

Tobias Klug, Thomas Ziegert, Andreas Zinnen, Victoria Carlsson

SAP Research CEC Darmstadt

Abstract

This paper reports experiences with applying UCD methods obtained during the development of a wearable prototype for the ward round in a hospital. Considering the non-computer related primary task in the design of the wearable system introduces a number of issues to applying existing UCD methods. This paper identifies a number of challenges that arise using specific UCD methods, as well as challenges that cut across the whole UCD process. These challenges and their implications are briefly discussed. However, the main focus lies on uncovering the breadth of problems.

1 Introduction

Wearable computing has been an active research topic for quite a while. While many of the technological challenges have been solved, there is an increasing demand for research in the area of usability. Wearable computing considers scenarios where today's desktop and mobile solutions are inappropriate. Unlike well researched desktop or mobile scenarios, users in wearable computing are away from their desk and have to carry out complex tasks in the real world. According to Gorlenko and Merrick (Gorlenko & Merrick, 2003) we call this kind of scenario the field context. The primary task requires most of the user's attention that often cannot be redirected to the computer. For this reason end user involvement becomes very important. Therefore, applying UCD principles to the development process is a natural choice. However, current research in mobile usability (Kjeldskov & Stage, 2004) shows the insufficiency of the available usability engineer's tool set. With wearable computing scenarios the problems are magnified, because the user's primary task needs to be integrated into the interaction.

The question of how to successfully apply existing UCD methods in such settings is still an open issue. Practical research on this topic is rare. This paper focuses on the challenges and pitfalls when applying UCD to wearable applications. The reported experiences were gained during the development of a wearable prototype for the ward round in a hospital. In a first step, potential use cases for wearable technology were identified in cooperation with the hospital staff. User observations, interviews and focus group discussions revealed a scenario focusing on the ward round. Several UCD activities were carried out resulting in a usability test of the final prototype.

The remainder of this paper is structured as follows: section 2 investigates related work regarding the application of UCD in non-desktop scenarios. Section 4 describes some of the UCD activities applied and discusses problems encountered because of the field context setting. Section 5 focuses on issues independent of the methodology used. Section 6 finally draws conclusions and indicates necessary future work.

2 Related Work

Wearable computing research is still quite young and most research still concentrates on getting the technology and infrastructure right. There is not much research addressing the question how to adapt the software development processes to the particularities of the field context.

Terrenghi et al. investigate how the mobile context of use changes the requirements engineering process and which new issues need to be taken into account when user centered design is applied to a mobile scenario (Terrenghi et al. 2005). However, their focus is purely on mobile applications. No insights how UCD methods need to adapt to these new issues are provided. Gorlenko and Merrick (Gorlenko & Merrick 2003) describe the differences between desktop applications, mobile office context and field context. They argue that everyone involved in developing applications for the mobile office and field contexts must focus on the total user experience. They also describe why current UCD methods are not well prepared for developing for the field context. The paper discusses effects of the field context on observations, task analysis, prototypes, design evaluation and validation. These papers discuss UCD issues in theory. The works are not based on hands on experiences collected during design of such applications. In fact, most papers published in this field are of a rather theoretical nature.

However, there are some notable exceptions. Kangas et al. (Kangas & Kinnunen 2005) report experiences with UCD activities which were collected during development of two mobile phone applications. Although these applications belong to the mobile office context, the recommendations given are also valid for the field context. Relevant questions in both contexts are the need for real prototypes when testing new interaction styles and the question whether lab testing is feasible when the context of use might have a high impact on usability.

Kristoffersen et al. (Kristoffersen & Ljungberg 1999) investigate computer support for inspection and engineering in the field. One of their important findings is that the challenges of mobile computing aren’t solely a result of the limitations of devices, but also of the work situation at hand. This is because the user needs to “make place” for the mobile device if it is not specially designed for the situation. Their goal is to design mobile applications in a way that allows interaction to “take place” without interrupting the user’s primary task. This is also one of the most important aspects when designing for the field context.

3 Case Study – Ward round

During the ward round, doctors need bedside access to patient records. In a paperless hospital of the future, patient records may be displayed and searchable via a bedside display. Beyond accessing patients’ records, doctors need to record findings and setup appointments for further patient treatment. Both, the findings and the appointments have to be stored in and synchronized with the hospital’s information system immediately. The scenario as such does have several challenges one has to cope with:

- The several tasks the ward round is actually comprised of are done in collaboration between doctor(s) and nurse(s).
- Whenever getting into physical contact with the patient, doctors must disinfect their hands before they are allowed to operate any non-medical equipment again. To disinfect, the doctors need to interrupt their current task, walk to a dispenser mounted on the wall. After disinfecting, the hands remain wet for some time preventing immediate use of any equipment. Therefore, contact free interaction methods seem to be promising in order to save time.
- Task switching as such is a major challenge not just because of the fact that when using a computer our users have to switch between their primary task and the computer related task but also because task switching is necessary while doing the primary task itself (e.g. examining the patient, discussing with the patient and the nurse).

The prototype was developed as part of the wearIT@work project. wearIT@work is a European Union funded project running under the auspices of Framework Program 6, involving 36 partners from across Europe. The design and development followed an ISO 13407 compliant UCD process model. The different methodologies discussed in this section covered the first of four planned iterations of the process. The UCD process started with initial field observations, interviews and focus group discussions on site. The outcomes of these investigations were basic workflow descriptions, general usability requirements and initial design ideas. These ideas were evaluated in a subsequent study based on UI mock-ups. In line with the study’s results, requirements for a prototype were defined. Later, the requirements were discussed, evaluated and prioritized together with end users. Then a prototype was specified taking the most important requirements into account and using technology, available to the project (provided by partners working in various fields of hardware and software development). Finally, the prototype was set up on site and tested by a group of doctor/nurse pairs.

The effects of the field context on the various process steps varied. Therefore, the next sections describe only those process steps in more detail where interesting challenges specific to the field context were experienced.

3.1 Field Observations

Early in the project a series of observations was conducted, where doctors were shadowed during their daily ward round. The goal was to discover the activities performed during a ward round, together with information about the environment and its influences on the doctor's work. To integrate the system seamlessly with the current ward round process, quantitative data about exemplary instances of the current process was necessary to analyze typical work situations. Because activities during a ward round change frequently, a fine grained capturing of executed tasks within a range of seconds was necessary. This quantitative data also showed how often situations occur and how long they last.

During the observations, a number of challenges were encountered. First of all, the environment was completely unknown to the observer. When observing office users the basic setting is already known in advance and the observer has first hand experiences with this kind of settings. A ward round, however, is most likely new to an observer. As a consequence, it was difficult to assess the importance of events and environmental factors as well as to note the right things.

The speed at which unknown activities take place is another challenge the observer has to cope with. The ward round is a highly optimized and parallelized process. For example, a doctor might read the patient's file, order an examination, talk to the patient and document his visit, all in less than a minute. In a typical approach, one would interrupt the doctor and ask for clarification when things went too fast and were not understood. However this approach is prohibited by the very environment that is studied. The doctor is not performing the ward round in his office, but in the hospital based on a strict schedule. In the present example, the hospital administration did not approve any technique that would have interrupted the ward round or would have made the patient aware of the fact that a study was in progress. For this reason and also for privacy reasons, audio and video recordings were not allowed.

Jotting field notes on paper worked for the initial observations when trying to uncover the breadth of activities and environmental influences. However, capturing detailed temporal data during the observation was only possible with computer support (Klug 2007).

3.2 Mock-up evaluation

Based on observations and interviews, first design ideas were generated. Those ideas were evaluated using a mock-up system. The system allowed the doctor to view patient records on a Tablet PC and gave the nurse the opportunity to complete x-ray request on a PDA. The studies, which were done using the mock-up, pursued three goals:

1. Evaluate the feasibility of using the selected devices during the ward round.
2. Analyze whether the distribution of tasks between doctor and nurse was practical.
3. Compare speech and pen input as suitable input modalities for the Tablet PC.

All evaluation results were purely qualitative, no quantitative data was collected. The mock-up system simulated two tasks: browsing the patient's record and ordering a new examination. The idea was to guide the doctors and nurses through a few scenarios. To perform the studies in a realistic environment, a normal patient room with bed, chairs and tables was used. The UIs on Tablet PC and PDA were implemented using Macromedia Flash. Whereas pen input was already part of the Tablet PC; speech input was simulated using the Wizard of Oz approach.

For the sake of simplicity only some navigational paths through the dialogues of the user interface were implemented. Therefore, it was important that only available functions were chosen by the participants. Invented patient cases were supposed to lead the doctors through the UI of the mock-ups. In case of a problem, the participants were given hints by the examiners.

According to Hagen (Hagen et al. 2005), enactments entail role-play and enacted scenes, allowing the user to interact freely with the system in the right context. Usually, and in contrast to simulations, no tasks are predefined. Therefore, the approach chosen represents a mixture between enactment and simulation.

The doctors and nurses mostly followed the implemented paths of the system. Anyway, sometimes the implemented paths spoke against a doctor's 'intuition', such that s/he was reluctant to continue. This led to long discussions and the tasks were not completed. In these cases, the workflow was interrupted and an analysis of the workflow‘s suitability became impossible.

Beyond the aforementioned problem, some influential factors of the ward round could not be considered. For example, the patient as the center of attention and a trigger for most actions could not be included in the study. Furthermore, the participants did not perform their tasks under time pressure common during the ward round.

Despite these limitations the setup included enough context to evoke feedback about practical aspects of the workflow and devices. The Tablet-PC was regarded too heavy by most participants. Additionally, the device interfered with the physical patient-doctor interaction because there was no place to put it down in a patient room where it would not have been contaminated. Lastly, the Bluetooth headset for speech input could not be used together with a doctor's stethoscope. These aspects would practically not have been discovered without doing tests within the considered context.

This findings show the importance of considering the context of use in the test setup. Even an incomplete setting can lead to important findings. Especially during initial evaluations where the focus is less on performance evaluations, every piece of context that can be easily reconstructed can be of value.

3.3 Usability Test

Based on the results of the mock-up evaluation and requirements gained through further focus group discussions, a first prototype was designed and built. There were several reasons why the Tablet PC wasn't the best solution for enabling access to patient data. First, the Tablet PC is not comfortable enough, because using it disrupts the usual workflow. Second, disinfection is required between getting in contact with the patient and the device. The Tablet PC was therefore replaced by a display attached to the patient's bed and gesture interaction. An interaction wristband was used to recognize arm gestures. Those gestures were used to navigate the application.

The goal of the usability test was to evaluate how well task switches, i.e. examining the patient and interacting with the system work and how the chosen gesture interaction fits with the ward round setting. Another goal was again to evaluate the suitability of the collaborative workflow between nurse and doctor. As the gesture recognition technology was fairly new for our user group, an initial study with university students was conducted. The initial study identified technical and practical issues with the gesture recognition system and helped us to increase the reliability and usability of it.

In terms of physical context, the prototype experiment was set up like the mock-up experiment. However, a dummy - normally used for educational purposes - was added to play the role of the patient as shown in Figure 1. The dummy enabled a more realistic performance of physical examination tasks, which was necessary to achieve the goals of the study.

Figure 1: Test setup for the usability test showing the bedside display and the dummy.

Although the doctor's new user interface was fully functional, a simulation rather than enactments was used because of the experiences in the mock-up evaluation. The doctor was supposed to view several documents and examine the patient once in between. This was achieved by instructions contained in the documents which simulated the patient records. Consequently, there was little room left for participants to deviate from the planned sequence of actions. All relevant interaction steps were performed by the participants, even though the content of the documents and scenarios was not realistic. However some doctors performed the physical examinations a bit nonchalantly and not at all to the same extent as they would have done it, had there been real patients. Consequently, the influence of the examination task on the gesture interaction was unrealistically small. Therefore, the overall interaction paradigm could be evaluated; switching between physical examination and computer task could not.

Using a simulation instead of an enactment turned out to be the right choice. Because the task sequence of browsing arbitrary documents was not too realistic in a medical context, the doctors had no problems doing the “wrong” thing.

4 General issues using UCD

The issues described in the previous section were a direct result of the particular method which was applied in this specific context. We encountered also some general issues directly resulting from using UCD in the field context. These general issues had an impact on all activities carried out and are likely to have impact on other methodologies as well.

4.1 Lab vs. field testing

When evaluating computer systems developed for the field context, there are two basic approaches: testing in the lab and testing in the field. Testing in the field means evaluating the system in its target environment, e.g. testing GPS navigation software while driving a real car. Lab testing on the other hand means reproducing parts of the context of use in a controlled environment. The selection of the appropriate method depends on the goals of the study and on the context of use itself.

Field tests ensure a realistic primary task as well as physical and social contexts. A successful field test allows the most precise predictions about the usability of the final product. However, field testing has also drawbacks. Even though providing the most realistic context, field evaluations are resource intensive. There are also many variables involved affecting the interaction with the system, which can be difficult to control. Therefore, if usability problems are detected, it might be difficult to discover their cause. Another issue is the amount of functionality that needs to be implemented before a field test can be done. In the ward round scenario, a field test is impossible without access to the hospital’s patient data.

Lab testing, on the other hand, can be done with mock-ups only. When evaluating in a laboratory context, it is also possible to control all the variables. Here the examiner needs to choose the aspects of the context of use, he wants to include in the study. Because the context of use is only approximate, the quality of usability and performance predictions depends largely on that approximation. If important aspects are missing, the results might be completely invalid. Another challenge is to ensure that context simulations in the laboratory are as realistic as possible, especially when aiming to simulate a primary task, like the examination of a patient.

A reasonable amount of research has been performed where the physical context of use was simulated in the laboratory (Kjeldskov & Stage 2004; Baillie & Schatz 2005; Been-Lirn Duh et al. 2006). However, there are not many results available regarding the simulation of the primary task (Witt & Drugge 2006).

In our case study for both, the mockup evaluation and the usability tests, a laboratory context was chosen for several reasons. Firstly, the prototype was not fully functional in the sense of being integrated with the hospital systems. Thus, no real patient data could have been accessed and no real examination orders could have been entered, which would be unpractical when dealing with real patients in the field. Secondly, most of the doctors had no prior experience with a similar application and especially not with gesture interaction. Thus, it was expected that a great deal of attention and time – which cannot be spared in a real ward round - would be required for interacting with the system. Thirdly, the involvement of real patients in the experiment was not possible due to privacy reasons. Fourthly, it would not have been possible to control which and how often primary tasks were performed. Thus, it would have been difficult to make sure that the important interdependencies between human computer interaction and the primary task were actually present and that these could be individually accounted for. Of course, some of the factors present in the reality like time-pressure or natural communication with patients or colleagues could not easily be simulated.

4.2 Involving the Non-User

Moving away from desktop scenarios, computers will be spread everywhere. Herstad (Herstad et al. 2000) introduces the non-user as “a person, or group, that is not directly using the technology in question, but that at the same time is affected in some way by the use of technology”. The authors motivate non-user involvement in the design process. In the healthcare scenario, doctors and nurses are the end users of the introduced system, and wearable computers will influence the doctor’s surroundings. Patients are a major part of these surroundings. Motivated by a strong demand for a better patient treatment in the healthcare sector, the patient’s well-being is very important. Privacy issues restrain contact to patients and make patient’s inclusion in the design phase impossible. Instead, the patient’s view is mainly based upon doctors’ and nurses’ comments or the designer’s appraisal what makes it impossible to be evaluated.

The doctors’ and nurses’ comments regarding the patients’ view resulted in two assumptions affecting the proposed interaction methods for the ward round prototype system. First, patients will be annoyed by speech interaction. They will feel addressed if doctors use speech commands to navigate through the system. Second, they will be bothered by obtrusive navigation gestures. In other healthcare scenarios, e.g. when patients will be anaesthetized, speech interaction might be an adequate interaction solution. Nurses clearly know the processes of a ward round, therefore they will not be distracted by the system as non-users. Interviews and focus group discussions revealed another interesting effect. Doctors were worried about how using gestures might affect their relationship with the patient. They were concerned to make themselves look strange performing gestures. Involving the non-user is therefore quite important to eliminate any wrong conclusions about any impact of users on non-users and vice versa.

4.3 Communicating HCI findings

Another challenge of designing for the field context is communicating the scenario and related requirements to designers, developers and colleagues. Usually these team members have not experienced the scenario at a first hand, but still need to get a feeling for the scenario. When developing for an office context this is less of a problem. The office context is well known. Team members can mostly rely on their own experiences when designing or developing. The field context differs strongly from the office context in terms of the primary tasks and the environment. Discovering these differences is exactly the goal of many user centered activities. However, only a small number of team members can be directly involved with the end users. The rest of the team needs to be informed about the scenario through presentations and written reports.

The described case study proved that it is difficult to capture all aspects of the ward round in a written report. No matter how detailed the descriptions were there was still room for misunderstandings. When these misunderstandings are not discovered or the HCI expert is not available for discussions, this can lead to serious design problems.

Existing UCD methodologies like contextual design (Beyer & Holtzblatt 1997) try to capture information about the work environment and the sequencing of activities in graphical models to ease the understanding of the context. However, they offer only simple graphical notations. The ward round for example does not consist of strict sequences, but parallel activities that cannot be represented in sequence charts. Also information about modalities used for these activities is not captured.

The use of timing diagrams of concrete scenario instances for communicating temporal aspects of a primary task, as proposed by Klug (Klug 2007), is currently being investigated. Diagrams like these can be captured during observations as described earlier.

4.4 Conclusions and Future Work

This paper identifies challenges and pitfalls one has to cope with when applying UCD to scenarios in the field context based on a case study.

The field context adds new challenges to stationary and mobile scenarios. It requires the application of UCD as a must while leading to problems with established UCD tools and methods. Field observations, usually delivering mostly qualitative data about the scenario have to be extended in order to collect quantitative data and information, to allow an analysis of how a computer task can be performed in parallel. The decision whether to apply lab or field tests becomes more difficult when designing for the field context, because of the increased importance of the user’s context and primary task. Whether to perform tests using simulation or enactment turned out to be another question. For the case study described here, simulation was given priority in order to achieve a higher level of control enabling the testing of particular functionalities of the mock-ups/prototypes. Further research is clearly necessary to adapt UCD methods and tools to the field context.

5 References

Kjeldskov, J., and Stage, J. New techniques for Evaluation of Mobile Systems. International Journal of Human-Computer Studies 60 (2004), 599–620

Bardram, J.: Applications of Context-Aware Computing in Hospital Work - Examples and Design Principles. In Proceedings of SAC 2004: The Nineteenth Symposium on Applied Computing, Nicosia, Cyprus, March (2004)

Bardram, J.: Hospitals of the future – Ubiquitous Computing support for Medical Work in Hospitals. Proceedings of UbiComp 2003: The Fifth International Conference on Ubiquitous Computing, Seattle, USA, October (2003)

Noma H., Ohmura A., Kuwahara N., Kogure K.: Wearable Sensors for Auto-Event-Recording on Medical Nursing. In Proceedings of ISWC (2004): The eighth annual IEEE International Symposium on Wearable Computing, Arlington, USA, October (2004)

Gorlenko, L., Merrick, R.: No wires attached: Usability challenges in the connected mobile world. IBM Systems Journal Vol.42 (2003) 639–651

Terrenghi, L., Kronen, M. and Valle, C.: Usability Requirements for Mobile Service Scenarios. Proc. of Human Computer Interaction (2005)

Kangas, E., Kinnunen, T.: Applying user-centered design to mobile application development. Communications of the ACM Vol.48 Nr.7 (2005) 55–59

Kristoffersen, S., Ljungberg, F.: “Making place” to make IT work: empirical explorations of HCI for mobile CSCW. Proc. of the international ACM SIGGROUP conference on Supporting group work (1999) 276–285

Hagen, P. and Robertson, T. and Kan, M. and Sadler, K.: Emerging research methods for understanding mobile technology use. Proc. of the 19th conference of the computer-human interaction special interest group (CHISIG) of Australia (2005) 1–10

Baillie, L., and Schatz, R. Exploring multimodality in the laboratory and the field. Proc. IMCI (2005) 100–107

Been-Lirn Duh, H., Tan. G.C.B. Usability evaluation for mobile device: A comparison of laboratory and field tests. Proc. MobileHCI (2006) 181–186

Witt, H., Drugge, M.: HotWire: An apparatus for Simulating Primary Tasks in Wearable Computing. ACM International Conference on Human Factors in Computing Systems (CHI), Montréal, Canada, April 22-27, 2006

Herstad, J., Stuedahl, D., van Thanh, D.: Non-user Centered Design of Personal Mobile Technologies. In Proceedings of IRIS 23, Uddevalla, Sweden, (2000)

Beyer, H. and Holtzblatt, K.: Contextual Design: Defining Customer-Centered Systems. Morgan Kaufmann (1997)

Klug, T. 2007. Computer aided observations of complex mobile situations. In CHI '07 Extended Abstracts on Human Factors in Computing Systems (San Jose, CA, USA, April 28 - May 03, 2007). CHI '07. ACM Press, New York, NY, 2507-2512.