Testing the usability of well scaled mobile maps for consumers

Corné P.J.M. van Elzakker[1], Peter J.M. van Oosterom[2] & Ioannis Delikostidis[1]

[1] International Institute for Geo-Information Science and Earth Observation (ITC), Enschede, The Netherlands
[2] Delft University of Technology, The Netherlands

elzakker@itc.nl / p.j.m.vanoosterom@tudelft.nl / delikostid13285@itc.nl

Introduction

The use of geo-information in mobile devices (smartphones, PDA's) is constantly increasing. Map displays play a prominent part in this, although they suffer from limitations set by the small screens, local (outdated) map copies, storage capacity and processing power of these devices. With the availability of high bandwidth wireless connections it should be possible to overcome some of these limitations and up-to-date map displays may be generated at the right level of detail and adjusted to the needs of the users. In this context, a Dutch research project started in 2006 with User-Centered Design (UCD), generalization and mobile geo-applications as keywords. This paper starts with a short description of the nature and objectives of this research project and its achievements so far. Generalization is particularly relevant for mobile geo-applications. Not only because zooming is a very important way of interaction of users with the map interface, but also because the progressive transfer of geo-information from server to device may be beneficial to usability (e.g. first retrieve 'rough' representations and get additional detail when required). With smooth zooming, realized through progressive transfer, map scale may change continuously in an animation style of visualization. This may help users to orient themselves. Part of the research deals with the application of the tGAP (topological generalized area portioning) structure in variable scale data structures suitable for progressive data transfer. The second section of this paper elaborates on this.

These generalization concepts will be implemented in a prototype of a mobile geo-application that will be developed through UCD techniques. The usability of the prototype will be evaluated in several iterative steps. The main objective of this paper is to present a methodology that may be used for this evaluation. First, a brief overview will be given of some research methods and techniques that have already been used in testing the usability of existing mobile geo-applications as well as those that have not been used before but demonstrate great potential. An analysis of these methods resulted in several test scenarios which have been put into practice with a mobile geo-application that is comparable to the prototype that is under development in the main research project. The outcome is a proposed methodology for the usability evaluation of mobile geo-applications, and particularly the interactive map displays thereof.

UWSM2: the research project on usable mobile maps

The research project on *Usable (and well scaled) mobile maps for consumers* [UWSM2, URL1] is executed by a consortium of research & development organizations (Delft University of Technology, ITC and TNO Defense, Security & Safety), software companies (ESRI and 1Spatial) and end-user organizations (Municipality of Amsterdam and ANWB, the Dutch Automobile Association). The project is partly funded by the Space for Geo-Information innovation programme [RGI, URL2] of the Netherlands' government. The overall objective of the project is to find solutions for multiple / vario-scale representations of geographic information and interaction with the consumers via mobile devices with context-aware, easy to use interfaces in a connected (wireless Internet) client-server setting. As such, two scientific challenges are addressed: automatic generalization and the human factors aspect of mobile applications. The challenge is to dynamically tailor the human-computer interaction to the user and momentary usage context. For example: by allowing quick access to specific details and supporting adequate browsing behaviour when needed.

The whole research project is driven by sample cases suggested by the municipality and the ANWB, involving user questions like: Where can I park my car and what will it cost? Where is the museum? How can I get there? and How can I get back to my car? These were combined into a *scenario*: a verbal description of possible uses of the system to be designed. The scenario describes the user and his/her tasks, including his/her goals, desires, and the context of use. It was the starting-point for further specification of the desired functionality through use case and interaction modeling with the help of UML, the Unified Modeling Language. As such, use case models describe frequently executed or complex tasks. The use cases may be related to each other, or be an aggregate of "smaller" use cases. For example, the use case "store location of car" may be part of the "bigger" use case "go to point of interest (POI)".

UWSM2 is organized into a number of Work Packages (phases), following a sound, iterative user-centred design methodology (see e.g. Van Elzakker & Wealands, 2006). The use case and interaction modeling was the outcome of the first Work Package (WP1). The results are now used for the development of a working prototype (WP2). The usability of the first prototype will be evaluated in WP3 and the results of this first testing will be used to improve the prototype in WP4. A final usability evaluation will be carried out in WP5 that should be completed by the end of 2008. This paper proposes a methodology for the usability evaluation. However, before we will address that issue, we will briefly deal with the generalization aspect of this research project in the next section.

Generalization for mobile geo-applications

Due to small screens of the mobile devices, which cannot show much of the map, a user needs to pan and zoom a lot to get a sufficient spatial understanding. That is, a feeling of the sizes, directions and distances between the relevant objects and their context. However, after a zoom or pan action, in nearly all mobile geo-applications a complete redraw is performed. In that case, the user often looses the 'mental' contact between the two maps. Initial experiences show that users are getting lost and do not build up a good mental map in order to support their task. Therefore, they may not appreciate new applications based on mobile maps. One of the key solutions to the described problem of user disorientation is vario-scale maps. Data structures supporting variable scale data sets are still very rare. There are a number of data structures available for multi-scale databases based on multiple representations (MRDB's), i.e. data to be used for a fixed number of scale (or resolution) intervals. These multiple representation data structures try to explicitly relate the corresponding objects at the different scale levels, in order to offer consistency during the use of the data. Drawbacks of the multiple representation data structures are that they do store redundant data (same coordinates, originating from the same source) and that they support only a limited number of scales. Another drawback of the multiple representation data structures is that they are not suitable for progressive data transfer, as each scale interval requires its own (independent) graphic representation to be transferred.

The tGAP structure is a data structure supporting vario-scale vector data. In earlier research both the theoretical and practical (implementation) aspects of the tGAP structure (topological Generalized Area Partitioning) have been described (van Oosterom, 2005; van Oosterom, de Vries, Meijers, 2006). Purpose of this tGAP structure is to store the data only once, with no redundancy of the geometry, and derive different representations of the same data on the fly according to the level of detail required. The tGAP structure can be used in a number of different ways: 1. to produce a representation at an arbitrary scale (a single map), 2. to produce a representation with feature(s) of interest at a larger scale and the surrounding features at smaller scales (a non-uniform scaled map) or 3. to produce a continuous range from rough to detailed representations. These three ways of using the tGAP structure are all useful. However, the smooth zooming, realized through progressive transfer does seem to be the most promising solution for mobile maps. Currently, research is conducted to improve the tGAP structure with respect to the theoretical, functional, and practical aspects. But in the end, the results should be usable as well.

Methods and techniques for usability testing of mobile geo-applications

There are two basic usability testing methodologies for the determination of the usability of mobile geo-applications: laboratory-based and field-based. They each have their advantages and disadvantages depending on the functions that have to be assessed and the context of use. Considering the special nature of mobile geo-applications, where the user interacts at the same time with the device / application and the natural environment, usability cannot be properly checked by means of controlled laboratory experiments alone. In the laboratory, a big part of the contextual information cannot be investigated and real users' behaviour and activities may not be sufficiently understood (Mennecke & Strader, 2003; Leitner et al., 2006; Kaikokken et al., 2005). Despite this, most studies on the usability of mobile geo-applications are executed in the laboratory, while only 19% is done in the field (Kjeildskov & Graham, 2003; Kaikokken et al., 2005).

For our research, we investigated some characteristic examples of usability testing of mobile geo- applications: the GiMoDig project (Nivala & Sarjakoski, 2005; Sarjakoski & Sarjakoski, 2005), the TIP system (Hinze & Buchanan, 2005) and the Trammate project (Kjeldskov et al., 2005). An UCD approach was followed in the GiMoDig project, involving the assessment of context-aware mobile map-based prototypes. Heuristic and expert evaluations were among the testing methods. Observation techniques were used in the TIP project, but no reference is made to user testing in the real context of use. The actual motivation behind the Trammate project was a comparison between different usability testing and data collection methods, executed both in the field and in the laboratory. Among the evaluation methods used were heuristic walkthrough, rapid reflection, thinking aloud, video and audio recording, observation and data logging.

The conclusions are that each of the methods and techniques applied had the ability to exclusively address specific usability problems; consequently, usually a combination of them should give the best results. Besides, a combination of laboratory and field testing seems to be the best solution as it comes to deep investigation of different usability problems that cannot be found with each approach alone. Laboratory-based testing can thoroughly assess interface and representation usability, whereas field-based testing can uncover user-application interaction in real / natural mobile environments. Another conclusion was that usability evaluation of mobile geo-applications requires a high amount of human resources.

Methods and techniques applied and put to the test

In the UWSM2 project, where satisfactory scalability and usability of mobile maps in real contexts of use is the main object of research, a combination of laboratory- and field-based testing should be performed. In view of the fact that there is not much experience yet with field-based usability evaluation we decided to pay special attention to the usability evaluation methodology to be used in the field and to the testing of several alternative combinations of research techniques.

Therefore, first of all we explored the use case and interaction models and the future prototype functions and we have extracted several user tasks from them. These were the understanding of the geographic location of the user through the mobile map, searching for the required POIs, receiving online location-based information, the presentation of the calculated route to the POIs, and obtaining navigation assistance through textual and graphical information. Converting these user tasks into questions that the system is aimed at answering, in this way supporting the decision making procedure of the user, was the next step. For the sake of the UWSM2 project, we focused on questions related to map scaling features, so that we could assess the clearness, consistency and comprehensibility of the map representation in different scales, the user's easiness of understanding his/her position on the map in relation with the real surroundings in all zoom levels and the helpfulness of the zooming/panning/rotation/orientation functions for better understanding of the map/route (Delikostidis, 2007).

Efficiency, effectiveness and user satisfaction as the fundamental attributes of usability were assessed through several qualitative testing methods and techniques in field-based user surveys. The methods selected were observation, thinking aloud, video/audio recording and semi-structured interviews. Although all of them have been used in the past for the evaluation of mobile as well as other applications, we thought there was ground for improvement so that they could better fit into a field-based testing environment, benefiting from the use of advanced technical solutions. Three different combinations of research techniques were created, which were later applied in 3 test groups of users (see Fig.1).

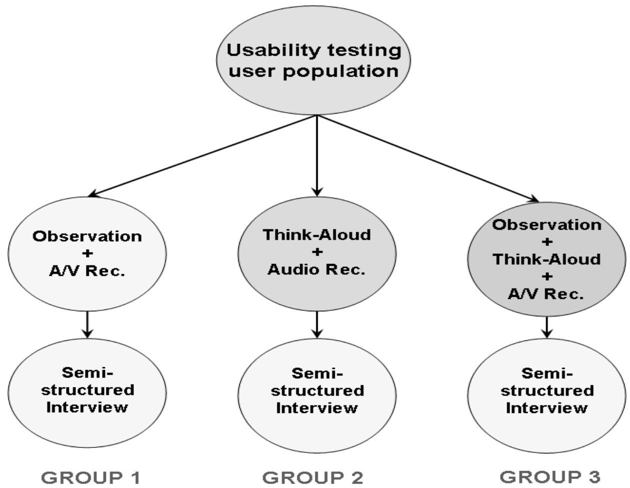

Figure 1: The 3 different combinations of techniques applied in 3 test user groups

Two main issues were considered during the design of the research experiment. The first was the possible influence of the researcher (observer) on the behaviour of the subject during the test and the second the possible need for more than one researcher to be involved in the test execution. The solution was the technical design and construction of a special field survey system that would enable one researcher alone to carry out the whole testing. We were looking for a solution better than just video recording through a camcorder, as the analysis of the test data would be cumbersome with the latter solution. Besides, the researcher would have to be dedicated to only use the camcorder, continuously aligning it in order to have a good view of the mobile device display,

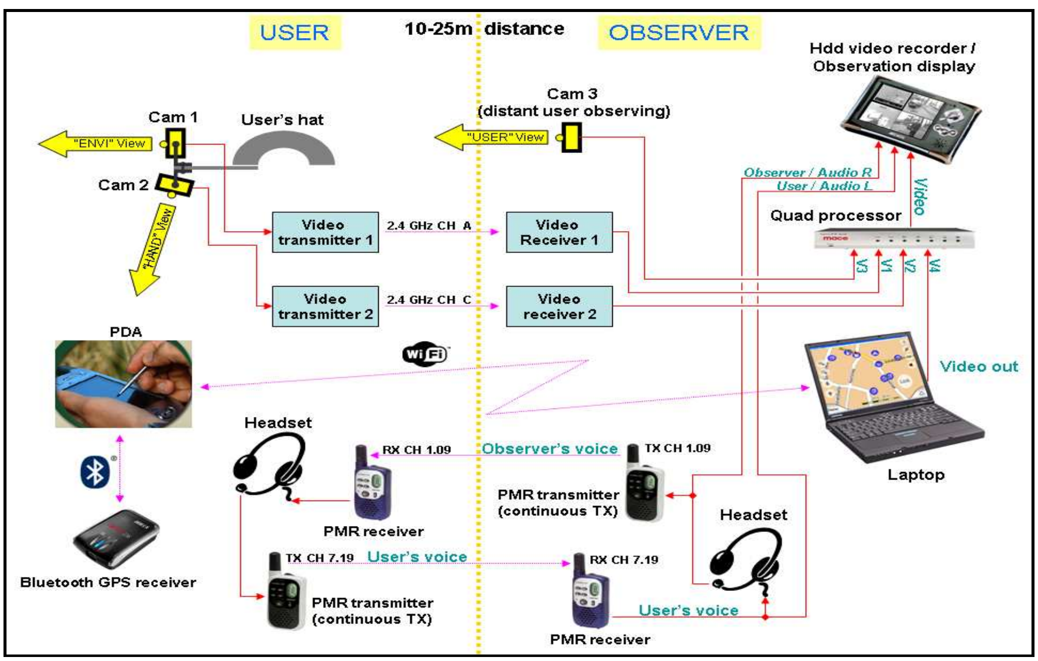

Figure 2: Overview of the implemented field observation system

alternating with views of the environment and the subject. Most likely, this would probably also pose a significant bias, as the observer would have to stay very close to the user during testing.

The system implemented enables the observer to stay at a fair distance of 10 to 25 meters from the test person, remotely observing, capturing and communicating with him/her during the test sessions. The system consists of three B&W wide angle mini cameras, two pairs of video transceivers, a video quad processor, a handheld video/audio recorder, a laptop, two pairs of modified PMR audio transceivers, a Wi-Fi and Bluetooth-enabled PDA, a Bluetooth GPS receiver and several lithium-ion and Ni-MH batteries (Fig. 2).

Two of the cameras were installed on a hat that the user wore, capturing the environment in front of him/her and his/her hands interacting with the mobile device (a HP iPAQ hx4700 PDA on which the mobile geo-application was running). The video signals from the two cameras were transmitted wirelessly to the observer through the pair of video transceivers and another video signal was obtained from the third camera that was carried by the observer on his chest and was capturing the user from a distance. The laptop with video output capability, carried by the observer in a backpack, was remotely capturing the screen of the PDA through a Wi-Fi connection in near real-time. The three video signals from the cameras together with the PDA screen capture through the video output of the laptop were connected to the quad processor and the output of the quad was connected to the handheld audio/video recorder. The output of the quad processor is a cross-like frame including all four video signals in one screen (Fig. 3).

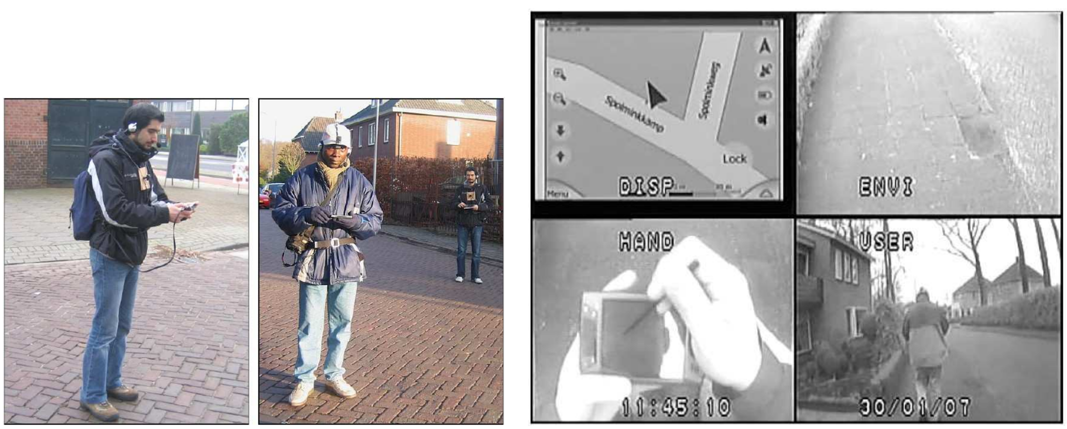

Figure 3: The observer/researcher and the test person with cameras on his hat and an actual screenshot of the quad processor video output

The PMR audio transceivers were used for the remote communication of the observer with the user and for the audio recording of their thinking aloud through the handheld audio/video recorder. Being modified, these transceivers allowed the continuous full-duplex communication between the observer and the user without the need of pressing the transmitting button of the corresponding transceivers.

Testing participants			**Knowledge / experience**				
Group 1	**Group 2**	**Group 3**	**GPS**	**Paper maps**	**Digital maps**	**PDA**	**Navigation**
A1	B1	C1	++	++	++	NO	+
A2	B2	C2	+++	+++	+++	+++	+++
A3	B3	C3	++	+++	+++	NO	NO
A4	B4	C4	++	++	++	++	+
A5	B5	C5	++	++	++	++	NO
A6	B6	C6	+++	+++	+++	+++	+++

Figure 4: The composition of the 3 test groups

After the construction of the field survey system, the next step was the selection of test persons. As this part of the research project was executed in another part of the Netherlands, the target population of the UWSM2 prototype (citizens of Amsterdam and tourists to the city) could not be used. Considering that the main aim was the comparison of different usability testing methodologies, ITC students (coming from all parts of the World) were found to be a good alternative. From this pool of test persons who, like tourists, do not know the survey area, homogeneous and comparable user groups could easily be formed. The selection was made through pre-questionnaires, asking for age, gender and knowledge/experience in different fields such as paper and digital maps, GPS and navigation systems and mobile devices. Based on their answers three comparable user groups of 6 test persons each were formed, including four men and two women, aging from 25 to 40 (Fig. 4). It was also made sure that the test persons were unfamiliar with the survey area, the village of Lonneker.

In order to assess the three proposed usability testing methodologies for a prototype that was not there yet, an existing mobile geo-application with generally similar functionalities had to be selected. A series of criteria was created, assessing different aspects of the candidate applications, including a.o. smooth zooming functionalities, different amount of detail/information in different zoom levels, zooming/panning/rotation functionalities, availability to the researcher, and detailed coverage of the study area. Although UWSM2 aims at using advanced on-line map retrieval techniques and real-time generalization, in this part of the research we focused on the map scaling functionalities only, using an offline existing mobile geo-application. In the end, the iGO My way 2006 application was selected for our testing [URL3].

In setting up the tests, efforts were made to keep the dynamic context variables inside acceptable limits, including rules for the test execution times, amount of light, weather conditions and environmental noise. The test sessions were scenario-based, including 7 navigation tasks in a predefined order to corresponding POIs. The users, representing visitors to an unfamiliar city, were supposed to use the functions of zooming, panning, rotation and orientation in order to successfully navigate to those POIs.

Proposed methodology for testing the usability of well-scaled mobile maps

After the completion of the surveys and analysis of the results (as summarized in Fig. 5), it was found that the combination of observation, thinking aloud and semi-structured interviewing with a remote observing/recording/communication system, as applied with test group no. 3, appeared to be the most promising methodology to be used for the field based evaluation of UWSM2 project prototype. It investigates several usability problems, gives a lot of information about the user and allows for a deeper analysis of the results. At the same time, the influence on the subject is kept low, as the researcher remains relatively invisible to him/her during the tests. The methodology can be further extended with more advanced functions and techniques like automated analysis of the results, or alternative techniques such as post survey questionnaires, GPS data logging (time and location, implying speed, turning, slowing down, etc.) and remote PDA data logging.

Methods / Group 1		Methods / Group 2		Methods / Group 3	
Advantages	Disadvantages	Advantages	Disadvantages	Advantages	Disadvantages
Video recording allows visual investigation / confirmation	No insight in the user's thoughts during task execution	The easier to implement	No visual investigation / confirmation	audio/visual investigation / confirmation	The most complex
Observation allows for thorough investigation of user actions	Complexity not significantly lower than group 3	Investigation of user's thoughts	audio recordings alone make analysis more difficult	Observation allows for thorough investigation of user actions	Systems used need a thorough testing before applying
	Battery power an issue		Long battery power	Insight of user's thoughts	Battery power an issue

Figure 5: Comparison of the three usability evaluation methodologies

Conclusion

This paper presented the findings of a research project into a feasible field-based usability evaluation methodology for a mobile geo-application in which generalization, as a consequence of smooth zooming, plays a prominent role. The proposed methodology, including a new technical solution for field-based qualitative usability research, can be used to evaluate the prospective UWSM2 prototype but, no doubt, other mobile geo-applications as well.

URLs

URL 1: UWSM2, the research project's website *http://www.gdmc.nl/uwsm2/*

URL 2: RGI, Space for Geo-Information innovation programme *http://www.rgi.nl/?l=eng*

URL 3: iGO My way 2006 navigation software *http://www.i-go.com/en/*

REFERENCES

Delikostidis, I. (2007), Methods and techniques for field-based usability testing of mobile geo-applications, MSc thesis, International Institute for Geo-Information Science & Earth Observation (ITC) Enschede, the Netherlands. March 2007. Available online at: *www.itc.nl/library/Papers_2007/msc/gfm/delikostidis.pdf*

Hinze, A. & Buchanan, G. (2005), Context awareness in Mobile Tourist Information Systems: Challenges for User Interaction. In: International Workshop on Context in mobile HCI at the 7th International Symposium on Human Computer Interaction with Mobile Devices and Services (Mobile HCI 2005), Salzburg, Austria, 19 - 22 September 2005. accessed on 2.11.2006 from *http://www.cs.waikato.ac.nz/~hinze/isdb/publications/hinze_buchanan_CONTEXTmHCI.pdf*

Jordan P. W. et al. (1996), Usability evaluation in industry. London: Taylor & Francis LTD.

Kaikokken A. et al. (2005), Usability Testing of Mobile Applications: A comparison between Laboratory and Field Testing. In: Journal of Usability Studies, issue 1, vol. 1, pp. 4-16.

Kjeldskov, J. & Graham, C. (2003), A Review of Mobile HCI Research Methods. In: Proceedings of the 5th International Conference on Human-Computer Interaction with Mobile Devices and Services (Mobile HCI 2003), Udine, Italy, 8-11 September 2003. Berlin etc.: Springer-Verlag. Usability and HCI Research methods, pp. 317-335.

Kjeldskov et al. (2005), Evaluating the Usability of a Mobile Guide: The influence of Location, Participants and Resources. In: Behaviour and Information Technology (BIT), vol. 24, no.1, pp. 51-65.

Leitner, G., Plattner, S. & Hitz, M. (2006), Usability Evaluation of mobile Applications in leisureoriented Contexts. In: Proceedings of the 13th International Conference of Information and Communication Technologies in Tourism (ENTER 2006), Lausanne, Switzerland, 18 - 20 January 2006.

Mennecke B. E. & Strader T. J. (2003), Mobile Commerce: Technology, Theory, and Applications. London: IRM Press.

Nivala, A-M. & Sarjakoski L.T. (2005), Preventing Interruptions in Mobile Map by Personalisation. Department of Geoinformatics and Cartography, Finnish Geodetic Institute, Masala, Finland.

Sarjakoski, L. & Sarjakoski, T. (2005), A use case based mobile GI service with embedded map generalization. Finnish Geodetic Institute, Masala, Finland

Van Elzakker, C.P.J.M. & Wealands, K. (2006) Use and users of Multimedia Cartography. In: W. Cartwright; M.P. Peterson & G. Gartner (editors), Multimedia Cartography, Second Edition. Chapter 34, pp. 487-504. Berlin etc.: Springer.

Van Oosterom, P. (2005), Variable-scale Topological Data Structures Suitable for Progressive Data Transfer: The GAP-face Tree and GAP-edge Forest.Cartography and Geographic Information Science, vol. 32, no. 4, pp. 331-346.

Van Oosterom, P., de Vries, M. and Meijers, M. (2006), Vario-scale data server in a web service context. In the Workshop of the ICA Commission on Map Generalisation and Multiple Representation – June 25th 2006.

Williams, J. R. (2004), Developing Performance Support for Computer Systems: A Strategy for Maximizing Usability and Learnability. Boca Raton: Taylor & Francis CRC Press.