PERVASIVE APPLICATION EVALUATION WITHIN VIRTUAL ENVIRONMENTS

Vassilis-Javed Khan, Koos Nuijten, Nils Deslé
NHTV Breda Univesity of Applied Sciences
khan.j@nhtv.nl, nuijten.k@nhtv.nl, desle.n@nhtv.nl

Keywords: augmented virtuality, pervasive computing, pervasive advertisement, pervasive computing evaluation.

Abstract: In this position paper, we present a setup for evaluating users' experience of pervasive applications within a virtual environment. We review existing literature on mixed reality and pervasive application evaluation. A conclusion of that review is the potential of evaluating applications such as location-based services in a virtual environment. Finally, we present our plans of evaluating user experience factors of location-based advertisements in a virtual supermarket, highlight methodological considerations and sketch future research directions.

1 INTRODUCTION

1.1 Augmented virtuality & virtual environments

In their seminal paper Milgram and Kishino (1994) presented the concept of "virtual continuum" where real environments appear at the one end of the continuum and virtual environments on the other end. In between that continuum the concept of "Augmented Virtuality" (AV) refers to the merging of real world objects into virtual worlds. Since the time of that paper, AV has had an application in a variety of fields (Dubois et al., 2010) such as medicine, entertainment, game, architecture, just to name a few.

The widespread use of a virtual environment (VE) is obviously routed to its advantages. It is a completely controllable environment, easy to manipulate and use it for different purposes. Thus, it is perfectly suited for experimental research where researchers need to control the environment while at the same time be able to observe and meticulously record observations and participant responses (Dubois et al., 2010). Compared to a PC application or low-fi prototypes, a VE is able to reconstruct the environment and therefore the conducted research would offer participants the right stimuli to yield reliable results. Compared to laboratory settings a VE allows participants to reconstruct their experience of everyday settings while ensuring that those settings are exactly the same for all participants. Thus, a VE would score higher on ecologically validity when compared to laboratory studies.

However, developing a VE is still to some extent expensive and requires extensive technical skills to create. Besides, there would always be a difference with the real world and therefore validation studies would be necessary to measure the effect of that difference. The immersiveness of the VE can also have the disadvantage of creating a halo effect to participants. Thus, there is the chance that participants would respond differently to such an environment due to the fact that they would have not had experiences with a VE before, or as frequent as with the real world.

Despite the shortcomings, settings that would be impossible to control and test otherwise would be partly possible to control in a VE.

It is notable that although tried out for different objectives, more commercial settings such as of the retail world seem to have been overlooked. Nevertheless, there is a resent exception of Renner et al. (2010) in which the researchers used a virtual supermarket for evaluating interaction techniques. The case of supermarket research for product placement, advertisement and optimization of shopping experience is already established. A virtual supermarket would certainly facilitate experimentation of the aforementioned research objectives.

Furthermore, in the field of pervasive computing, applications are usually evaluated either in a lab or in-situ. Laboratory evaluations are probably very useful for usability studies. However, when it comes to evaluating more abstract concepts such as the user experience of a certain application a lab study would affect the ecological validity of the evaluation.

Using virtual environments to evaluate pervasive computing applications has been explored in the past. Ubiwise (Barton and Vijayaraghavan, 2002) is cited as one of the first attempts to evaluate pervasive applications with the assistance of a virtual environment. Since that first attempt there have been several other prototypes tested with the same purpose. For a review of those prototypes we would refer to Reynolds, et al. (2006) and Leichtenstern et al. (2010). There are two conclusions to be drawn from those research efforts. First, researchers have been targeting their efforts into either the technical development of testbeds of pervasive computing applications (Bruneau et al., 2009, O'Neill et al., 2005). Second, so far, to our knowledge, those efforts were simulating the VE on a desktop computer. The serious shortcoming of such a setup is the questionable level of immersion for the user (Leichtenstern et al., 2010). Thus, a VE presented on a desktop seems not to provide the expected benefits in early testing of pervasive computing applications.

In the next section we survey literature of in-situ pervasive application evaluation to analyse the pros and cons of such an approach.

1.2 Evaluation of pervasive computing applications

Current pervasive evaluation techniques include in-situ evaluations in which systems are being deployed in the real world. While trying to be ecologically valid drawbacks of this approach are considerable.

Issues that might seem trivial at first become serious obstacles for in-situ evaluations. In their distributed and heterogeneous system for supporting sergeants and hospital personnel Hansen et al. (2006) report of such an experience. To give an example, while the authors thought that finding a place to place their public displays would be trivial it actually proved to be a major problem. Trying to generalize their findings, the authors report no less than 19 items of a checklist to consider before actually deploying a prototype in the real setting. The authors categorized those items into three categories: hardware, software and user setting.

Items from all three categories could partly be addressed if a similar evaluation would take place on a virtual environment. For example, the security of the environment poses a threat to the equipment of the research team; the space that the prototype would need to use; the integration of the prototype with third-party systems; the developers' support and the organizational politics that might arise from such an evaluation could serve as examples which would be either obsolete or could be addressed in a virtual environment. Of course that would imply that the whole setting of a hospital, in this particular case, would have been implemented in a virtual setting. However, once it has been implemented, one could both run numerous studies and keep on enhancing that VE with more objects or actors.

Another issue that the authors raise that could again easily be addressed in a virtual environment is the number of users actually being able to experience the prototype that is evaluated. In a hospital situation, and especially in an operating theatre, it is understandable that only the designated doctors and personnel would have access to that. Thus, this is another practical issue that raises questions of reliability of the evaluation of such prototypes. What would happen with different doctors and hospital personnel can only be hypothesized in this situation. On virtual setting several users could test the same prototype.

In their evaluation of three pervasive applications Consolvo et al. (2007) also discuss the limitations of an in-situ evaluation. Their first study researched a prototype deployed in a home setting for supporting eldercare. The prototype was mimicking sensor data with a Wizard of Oz. In-situ data were collected daily over the phone. Their second study researched a location aware application to understand the link between user preferences and destinations. In this study, situ self-reports via messages triggered based on the participants' arrival at a destination were collected along with interview data. Their third study investigated daily physical activity and whether sharing activity related data with a small group of friends might influence physical activity goals. This study combined interviews and questionnaires with in situ user-initiated logging of pedometer data mobile phone application.

The authors report that "essential usage context and pragmatic environmental constraints that might otherwise go unnoticed in more controlled settings" are necessary to actually evaluate pervasive applications since the context of use changes and plays an important role in the user's motivation and experiences of the technology.

However, for the first study, the authors also report: “a researcher had to call the elder or primary caregiver up to six times per day from the early morning until late evening every day for four deployments, each of which lasted for 3 weeks”. This makes one critique how ecologically valid would such an evaluation be since the actual use of such an application would obviously differ a lot. Besides, the authors also report other shortcomings; participants had established a rapport with one particular researcher and preferred talking with that researcher only; the study was “fairly labour intensive” and “required careful planning, preparation, coordination, and effort”. One would certainly appreciate the difficulties in executing such a study but at the same time still wonder how much would the study itself distort the users' experience.

Khan et al. (2010) also report on the challenges of in-situ evaluation of a pervasive computing application targeting working parents. Apart from technical issues that were also reported in the aforementioned studies, trivial unforeseen problems such as participants having to carry a second mobile device next to their own phone also negatively affect the evaluation efforts.

Although the theoretical benefits of evaluating pervasive computing applications in-situ are substantial, the challenges are still considerable. In some cases while pursuing to overcome those challenges the ecological validity of evaluation studies becomes questionable.

A concluding remark is that the use of a VE has been limited when it comes to evaluation of pervasive computing applications; previous studies have asked users to experience a VE on a PC. We envision exploring the use of a VE, not presented in a PC but in an environment where the user is actually immersed, for pervasive application evaluation, since the context of usage can to a great extent be reconstructed.

2. CASE STUDY

Our interest lies in the evaluation of pervasive advertising in a setting of a virtual supermarket. Our lab has a CAVE (Cave Automated Virtual Environment) consisting of four display walls of 3.5 wide by 2.6m high. These displays form an enclosed room, and images are back-projected on them. The projections are calculated in such a way that, when a user stands in the room, the illusion of a continuous, 360 degree, view is created (Figure 1).

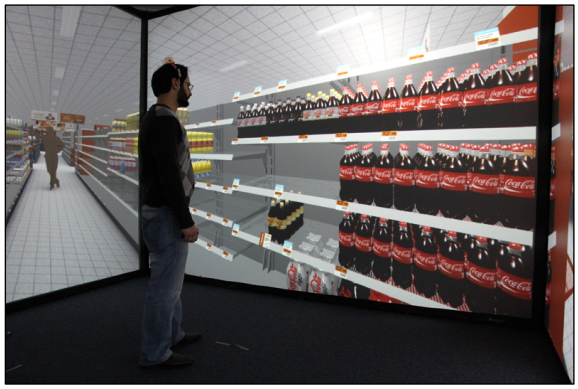

Figure 1: Participant interacting with the virtual supermarket. The participant is wearing a head tracker.

We have created a head-tracking technique using four Wii-motes, one in each corner of the room. The user while being inside the CAVE wears a headband with an infrared LED array. Through the 2D coordinates each Wii-mote returns to the server, we triangulate an exact 3D position of the user's head. We use this position to control motion in the CAVE, by using a "human joystick" scheme, whereby the virtual camera will strafe in the direction the user is standing, relative to the CAVE's center.

Virtual worlds are created in Maya, and rendered using OGRE as the underlying 3D engine. We have created several layers of frameworks to be able to do the network synchronisation, head tracking, model importing and optimized rendering.

For our first case study, a virtual supermarket has been created in Maya. The supermarket is a generic supermarket, based on common denominators found in typical Dutch supermarkets. It contains isles, freezers, checkout counters, etc., all rendered with realistic graphics. Currently, about 1/6 of the shelves are filled with products, but it is our near future plans to fill it with a representative product set.

2.1 User experience evaluation of location based advertisements

As a first case study we have planned an experiment to evaluate aspects of user experience of location-based ads within the virtual supermarket. It has been shown that perceived intrusiveness of ads leads to ad irritation (Li and Lee, 2002). High values of ad irritation lead to high a probability of ad avoidance (Li and Lee, 2002).

We are interested in exploring whether location affects the way people perceive ad intrusiveness. The hypothesis is that when people would receive an ad in their mobile phone close to a product that is

related to the ad, they would not perceive that ad as intrusive and would not be irritated by it. In other words, we hypothesize that when there is a match between an ad and the supermarket shelf in which the advertised product is sold people will find the ad less intrusive.

To test this hypothesis we developed an Android mobile phone application. The application interacts through Bluetooth with the virtual supermarket. When the user, carrying the mobile phone, is within a certain distance of a virtual shelve then the phone vibrates and sets off an alert sound while presenting an ad.

2.2 Long term plans

Going beyond the case study of perceived intrusiveness of location based advertisements, our goals extend into exploring a variety of pervasive computing applications in the virtual environment.

More generic questions include validation studies testing the perceived presence of the virtual supermarket with a real supermarket as well as exploring the added value of such an immersive virtual supermarket with a PC version of similar virtual supermarkets. It is safe to hypothesize that the immersive version of a virtual supermarket might well provoke behavior that would resemble more closely the behavior of a real supermarket.

3. CONCLUSIONS

Having a plethora of online, virtual supermarkets on the browser of people's PC is something that is nowadays a reality. With the development of technology, one would wonder what kind of behavior would people exhibit when encountered with a virtual, immersive supermarket compared to the PC-based.

We set out to execute a series of studies in an augmented virtual supermarket. We have identified that pervasive application evaluation with the assistance of mixed reality environments is still in its infancy. The advantages of a simulated environment lie in the fact that such an environment is fully controllable and adaptable to the researchers' needs. Evaluating pervasive computing in-situ has many advantages nevertheless it also poses various challenges. Many of the challenges could be tackled with the use of a virtual environment. It lies up to researchers to make use of the new possibilities that such a setup would create.

ACKNOWLEDGEMENTS

We would like to express our gratitude to our financial sponsors: the Dutch Ministry of Education, Culture and Science (OCW), RAAK scheme. Many thanks to Arief Ernst Hühn who took the photo displayed in this paper.

REFERENCES

Barton J.J., & Vijayaraghavan, V., 2002. Ubiwise: A ubiquitous wireless infrastructure simulation environment. *Tech. report HP*. HP Labs

Bruneau, J., Jouve, W., Consel, C., 2009. DiaSim: A parameterized simulator for pervasive computing applications. *In Proc. of MobiQuitous'09*, pp.1-10

Consolvo, S., Harrison, B., Smith, I., Chen, M. Y., Everitt, K., Froehlich, J., Landay, J., 2007. Conducting in situ evaluations for and with ubiquitous computing technologies. *IJHCI, 22*(1-2):103-118.

Dubois E., Gray, P., & Nigay, L. (eds.), 2010. *The Engineering of Mixed Reality Systems*, Human-Computer Interaction Series, doi:10.1007/978-1-84882-733-2_1

Hansen, T. R., Bardram, J. E. & Soegaard, M., 2006. Moving Out of the Lab: Deploying Pervasive Technologies in a Hospital. *IEEE Pervasive Computing, 5*(3):24–31

Khan, V.J., Markopoulos, P, Eggen, B., & Metaxas, G., 2010. Evaluation of a pervasive awareness system designed for busy parents. *Pervasive and Mobile Computing, 6*(5):537-558

Leichtenstern, K., André, E., & Rehm, M., 2010. Using the Hybrid Simulation for Early User Evaluations of Pervasive Interactions. *In Proc. of NordiCHI2010*, pp. 315-324.

Li Hairong, S. & Lee, J., 2002. Measuring the intrusiveness of advertisements: Scale development and validation. *Journal of Advertising, 31*(2):37–47.

Milgram, P. & Kishino, F., 1994. A taxonomy of mixed reality visual displays. *IEICE transactions on information systems, E77-D*(12).

O'Neill, E., Klepal, M., Lewis, D., O'Donnell, T., O'Sullivan, D., Pesch, D., 2005. A testbed for evaluating human interaction with ubiquitous computing environments. *In Proc. of Tridentcom2005*. pp. 60-69

Renner, P., Dankert, T., Schneider, D., Mattar, N. & Pfeiffer, T., 2010. Navigating and Selecting in the Virtual Supermarket: Review and Update of Classic Interaction Techniques. *In Virtuelle und Erweiterte Realität*: 7. Workshop der GI-Fachgruppe VR/AR, 71-82. Aachen, Germany: Shaker Verlag.

Reynolds, V., Cahill, V., & Senart, A., 2006. Requirements for an ubiquitous computing simulation and emulation environment. *In Proc. of InterSense '06*. Article 1. doi:10.1145/1142680.1142682