Virtual Museum Interface Evaluation: the ARCO case

Sylaiou S.[1, 2], Karoulis A.[2], White M.[1]

[1] University of Sussex, U.K.
[2] Aristotle University of Thessaloniki, Greece
sylaiou@photo.topo.auth.gr, athanasis@karoulis.gr, M.White@sussex.ac.uk

Abstract

The Augmented Representation of Cultural Objects (ARCO) system provides software and interface tools to museum curators to develop virtual museum exhibitions, as well as a virtual environment for museum visitors over the World Wide Web or in informative kiosks. The main purpose of the system is to offer an enhanced educative and entertaining experience to virtual museum visitors. In order to assess the usability of the system, two approaches have been employed: a questionnaire based survey and a Cognitive Walkthrough session. Both approaches employed expert evaluators, such as domain experts and usability experts. The result of this study show a fair performance of the followed approach, as regards the consumed time, financial and other resources, as a great deal of usability problems has been uncovered and many aspects of the system have been investigated. The knowledge gathered aims at creating a conceptual framework for diagnose usability problems in systems in the area of virtual Cultural Heritage.

Keywords: Usability evaluation, museum interface, augmented reality, cognitive walkthrough.

1. Introduction

Museum collections are the source from which the museum's unique role in the cultural fabric of society emanates via their contribution to scholarship, being the instruments of its education role, and the cause of its public enlightenment (Perrot, 1977). However, large collections and certain of the artefacts they hold, remain in storage places due to the museums' lack of space, the high cost of maintaining the exhibits and the fragility of certain cultural artefacts. Current research (Jones and Christal, 2002) and an extensive survey to European museum sector have shown (Tsapatori, 2003) that technologies such as the World Wide Web enhanced by 3D visualization tools can provide solutions to the aforementioned problems. In addition to these, the use and integration of the promising Virtual Reality (VR), Augmented Reality (AR) and Web3D technologies in conjunction with database technology may facilitate the preservation, dissemination and presentation of cultural artefacts in

museums' collections and also educate in an innovative and attractive way the wide public. Virtual Reality signifies a synthetic world, whereas Augmented Reality signifies computer generated 2D or 3D virtual worlds superimposed on the real world. Web3D is used to represent the application of XML (eXtended Markup Language) and VRML (Virtual Reality Markup Language) technologies to deliver interactive 3D virtual objects in 3D virtual museums (Liarokapis et al., 2004). Previous research has made use of 3D multimedia tools in order to record, reconstruct and visualize archaeological ruins using computer graphics (Cosmas et al., 2001) and also provides interactive AR guides for the visualization of cultural heritage sites information (Gleue & Dähne, 2001). Moreover, relevant research has demonstrated that 3D technology '*offers museums rich opportunities in a range of areas from public access to conservation*' (Shaw et al., 2004). In order to address these aspects, the ARCO (Augmented Representation of Cultural Objects) (ARCO, 2004) system has been developed and described in detail in Wojciechowski et al. (2004). In this paper we report on the usability evaluation of the two main components of the ARCO system, namely the ACMA (ARCO Content Management Application) and the ARIF (Augmented Reality InterFace) subsystems.

2. The ARCO System

The ARCO system allows museum curators to build, manage, archive and present virtual museum exhibitions based on 3D models of artifacts. ARCO also allows end-users to explore virtual exhibitions implemented using the system (Wojciechowski et al., 2004).

The cultural artifacts are digitized by means of a custom built stereo photogrammetry system (Object Modeler), mainly for digitizing small and medium sized objects and a custom modeling framework (Interactive Model Refinement and Rendering tool) that is used, in order to refine the digitized artifact (Patel et al., 2003). The 3D models are accompanied by images, texts, metadata information, sounds and movies. These virtual reconstructions (3D models and accompanying data sets) are represented as eXtensible Markup Language (XML) based data to allow interoperable exchange between ARCO and external heritage systems (Wojciechowski et al., 2004). These virtual reconstructions are stored in an Oracle9i database system and managed through the use of a specially designed ARCO Content Management Application, which also allows the museum to build and publish virtual exhibitions to the Internet or a museum kiosk system.

2. The ARCO components

Two main components of the ARCO system were of interest for the evaluation: the ARCO Content Management Application (ACMA) and Augmented Reality Interface (ARIF). ACMA allows publishing of virtual museums to both Web and a specially

designed application (ARIF) for switching between the Web and an AR system.

The ACMA application is implemented in Java and it includes the database of the representations of cultural objects and their associated media objects, such as images, 3D models, texts, movies, sounds and relevant metadata (Mourkoussis et al., 2003). It enables user-friendly management of different types of data stored in the ARCO database, through various managers, such as the *Cultural Object Manager* (deals with virtual representations of cultural artefacts), the *Presentation Manager* (manages virtual exhibitions with the help of templates) and the *Template Manager* (stores these visualization templates).

The ARIF component is a presentation or visualisation framework that consists of three main subcomponents:

- The *ARIF Exhibition Server*. Data stored in the ARCO Database is visualized on user interfaces via the ARIF Exhibition Server.
- The *ARIF Presentation Domains* with implemented web browser functionality, suited for web-based presentations.
- The *ARIF AR—Augmented reality functionality*. This sub-component provides an AR based virtual museum exhibition experience on a touch screen in the museum environment using table-top AR learning experiences, e.g. AR quizzes and on-line museum exhibitions.

3. Evaluation of the ARCO System

The ARCO System has been evaluated by utilizing a variety of methods, both empirical and expert-based, and some preliminary results have already been reported in Sylaiou et al. (2004). However, this study focuses only on the usability evaluation of the system and bases on two evaluation sessions, one questionnaire based and one session of Cognitive Walkthrough. The questionnaire based session was performed by the museum curators and assessed the ACMA as well the ARIF interface. The Cognitive Walkthrough was performed by "visitors" and concerned only the ARIF interface.

3.1. Participants

Ten domain experts took part in the evaluation aged between twenty-eight to sixty years old. All of them were museum curators from various departments of the Victoria and Albert Museum, London, UK. No end-users were involved in the technical development of the ARCO system, so they could not be employed to assess the ARCO interface. In contrary, the museum curators were involved in the technical development from an early stage setting user requirements and providing appropriate feedback during the early stages of implementation. So, they also have been employed as expert evaluators during this phase of the evaluation and have been

asked to fulfill the QUIS questionnaire.

In addition to this session, four students and two usability experts of the department of Informatics of the AUTh, Greece, acted as museum visitors and performed a Cognitive Walkthrough through the web-based ARIF interface, provided at http://www.arco-web.org/vmesite/V&A/VAMGallery.html. They were asked to assess the same aspects as the museum curators, namely the multimedia presentation in ARIF, however, they proceeded further and evaluated the overall usability under a technical point of view. These opinions are subsequently also presented.

3.2. Instrumentation

The QUIS (Questionnaire for User Interaction Satisfaction - Schneiderman & Plaisant, 2005) assessed museum curators' contentment while interacting with the ACMA and ARIF interfaces by means of a 9-scale Likert scale. This questionnaire was the main instrument to record their estimations. In contrary, the empirical evaluation used no questionnaire; however the same set of questions has been set to the usability evaluators. So, a direct comparison between the assessment of the curators (domain experts) and the "users-visitors" (usability experts) could be made.

The QUIS questionnaire consists of 7 parts. Part 1 concerns general experience with ICT (Information and Communication Technologies), and is not of great importance in this case, and has not been considered. Part 2 assesses the overall user reactions as regards to the evaluated system, Part 3 concerns the windows layout of the system, Part 4 the terminology used, Part 5 the learnability of the interface (how easy it is to learn), and Part 6 the system capabilities. These first 6 Parts evaluated the ACMA component, while the last Part 7 of the QUIS questionnaire concerned the multimedia presentation in ARIF, so, it could directly be combined with the evaluation of the usability experts in Greece, in order to elicit more accurate results.

It must be explicit be stated at this point, that this study concerns not the assessment of the value of the interface itself. There are a number of studies evaluating the ARCO system in a holistic manner, such as Sylaiou et al. (2004) and Sylaiou et al. (2006) with concrete suggestions for the improvement of the system. This study focuses merely on the comparison of the assessments of two different groups of expert evaluators, namely, the domain experts, who are aware of the cultural heritage domain, yet unaware of usability aspects, and the usability experts, who are aware of the usability aspects, yet can act only as users-visitors in a museum context.

4. Results

4.1. Session 1: Curators and QUIS

The museum curators (domain experts) evaluated by means of the QUIS a number of

aspects as regards both ACMA and ARIF interfaces. There were six groups of questions, evaluating the different parts of the system, as already described above.

4.2. Session 2: Cognitive Walkthrough in ARIF

The next session, performed at the multimedia laboratory of the Department of Informatics, AUTh., consisted of a Cognitive Walkthrough through the ARIF interface and, because no human artifact is perfect, pinpointed also a number of usability problems. However, the usability experts had a completely different view than museum curators and made some concrete statements, such as the resolution of the screen and the level of detail of the artifacts, which were more "puristic" than it could be depicted on the QUIS.

Accordingly, the usability experts have been asked to complete the same QUIS questions concerning the ARIF interface, namely only Part 7.

4.3. Interpretation and Discussion

As regards the first session, where only domain experts participated, the first obvious result of the statistics is the low mean value of almost all questions. The highest is at 6,56, and the lowest at 5,91. In a 9-scale Likert scale and given the relative high values people usually give in questionnaire based surveys, this is an indication of an overall "concerned acceptance" of the usability of the interface. In more detail:

The overall interface is considered to be powerful enough and flexible, although a little dull and frustrating. The handling of the various windows elements is assessed as most successful, providing the highest mean. However, the terminology used provided some scepticism. This is per se important, as museum curators are aware of the domain terminology; so this could be an indication that the terminology used in the system did not completely adhere to the domain standards. Furthermore, the system scaffolding was not adequate, as system messages, information on user progress or error messages were considered more frustrating than helpful. The learnability of the interface was also questioned, as well as the remembering of certain system commands. Finally, the multimedia presentation has been considered by the museum curators as sound.

This last point is however one of the most debatable of this study. During the second session, the usability experts in Greece, who acted as museum visitors and visited the museum through a web browser, reported a mediocre usability of the environment and a low satisfaction as regards to the cognitive aspects of the interface. They considered the environment to be unintuitive, without adequate help to scaffold novice users and with poor level of information as regards the presented artefacts.

The provided mean values are significant lower than those of the domain experts. So, a question arose here, namely whether the answers provided by the domain experts

are in accordance to those of the usability experts. In order to clarify this emerged aspect, a post-hoc elaboration procedure has been designed: An independent samples t-test as well as the non-parametric Mann-Whitney and Wilcoxon tests have been employed, in order to compare the mean values of the estimations of the two group of experts, as shown in Table 1 below.

Table 1. *a. Independent Samples t-Test*

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Part 7	Equal variances assumed	,917	,340	6,749	118	,000	2,08333	,30867	1,47208	2,69459
	Equal variances not assumed			6,589	92,420	,000	2,08333	,31618	1,45541	2,71126

b. Mann-Whitney and Wilcoxon tests(a)

	Part 7
Mann-Whitney U	689,500
Wilcoxon W	1865,500
Z	-5,638
Asymp. Sig. (2-tailed)	,000

In Table 1.a. the two rows represent the two possibilities, namely assuming or not the homogeneity of the variance of the samples. This homogeneity is depicted by means of the Levene's test, presented in the first two columns.

In all tests, the provided statistical significance of 0,000 (SPSS cuts the rest of the decimals, indicating a very small number) is presented in the column "Sig. (2-tailed)" for both cases and depicts a statistical significance at a level of a p-value lower than 0,001. So, *there is difference* of the evaluators' opinions due to the fact that they belong to different expert groups.

In Table 1.b. the row "Asymp. Sig. (2-tailed)" also provides a significance lower than p=0,001, so the non parametric tests MW and Wicloxon (more robust, although less accurate) result in the same conclusion.

Thus, the next point of discussion here is why there is such a great divergence between the curators and the usability experts' opinions. Some studies, such as Karoulis & Pombortsis (2000), report that usability experts are usually more rigorous than users. This explanation seems however in this case not plausible for two reasons. Firstly museum curators are also experts and should also be rigorous in their estimations, and secondly, the usability experts acted in this session as real users, who were initially enthusiastic to visit the virtual museum, yet they were at the end of the session not thus enthusiastic. The root of this problem is probably in the nature of the

evaluation. It is a fact that both groups have been asked to assess usability features of the interface. It is more plausible to believe that in this context, the usability experts are closer to the goal than the domain experts, as they know exactly what and how has to be investigated. It also seems a plausible claim the fact that museum curators have a more or less "foggy" impression of usability and its parameters, so in this context, errato humanum est…

5. Conclusion

The first obvious conclusion is that the usability evaluation of a museum virtual interface is possible through an expert-based approach. The museum curators are aware of many aspects on this domain and they perform adequately if they are surveyed in the correct way. Their responses lead to concrete improvements of the interface and their qualitative comments, in preliminary form already presented in Sylaiou et al. (2004) are a valuable source to improve such kind of interfaces.

However, the implication of the complete opposite thesis of the usability experts who acted as users raises many questions on some biases of the questionnaire-based surveys, as already stated. A tentative claim is that an expert-based usability evaluation cannot be performed without the participation of usability experts. However, in the described context, domain experts are also inevitable. As shown in the first 6 Parts of the evaluation, the museum curators showed an overall satisfaction on the usability of the system. The debate emerged when the usability experts considered the web-interface (the ARIF) acting as users. Here the "usability expert as a user" view was very different from the "domain expert as a curator" view.

So, final conclusion of this study is that one encounters here the limits of the expert-based interface evaluation approach: in complex interfaces, double experts (usability and domain experts) are inevitable for reliable and valid results. Simple experts (only domain or only usability) do not seem to perform adequately. However, the fact that such double experts are extremely rare and expensive, pinpoints the aforementioned limit of the expert-based approaches. This is of course a tentative claim, as this aspect was outside of the scope of the present study; therefore, new studies must be set up in order to validate these claims.

Acknowledgments

Part of this work has been funded by the Marie Curie Actions Human resources and Mobility Marie Curie training site: Virtual Reality and computer graphics, project HPMT-CT-2001-00326. Also, the authors would like to thank Katerina Mania for her suggestions during the first phase of the research.

References

ARCO, Augmented Representation of Cultural Objects (2004). http://www.arco-web.org/ (Retrived 7 Oct 2004).

Cosmas, J., Itegaki, T., Green, D. (2001). 3D MURALE: a multimedia system for archaeology. *Proceedings of the ACM-SIGGRAPH Conference on Virtual Reality, Archaeology and Cultural Heritage VAST 2001*, Athens, Greece, pp. 297-305.

Gleue, T., and Dähne, P. (2001). Design and Implementation of a Mobile Device for Outdoor Augmented Reality in the ARCHEOGUIDE Project. *Proceedings of the Virtual Reality, Archaeology, and Cultural Heritage International Symposium (VAST01)*, Athens, Greece.

Jones, J., and Christal, M., (2002). *The Future of Virtual Museums: On-Line, Immersive, 3D Environments*. Created Realities Group.

Karoulis, A., and Pombortsis, A. (2000). Evaluating the Usability of Multimedia Educational Software for Use in the Classroom Using a «Combinatory Evaluation» Approach. *Proc. of EDEN 4th Open Classroom Conference*, 20-21 Nov 2000, Barcelona, Spain.

Liarokapis, F., Mourkoussis, N., White, M., Darcy, J., Sifniotis, M., Petridis, P., Basu, A., and Lister, P.F. (2004). Web3D and Augmented Reality to support Engineering Education, *World Transactions on Engineering and Technology Education*, UICEE, 3 (1), pp. 11-14.

Mourkoussis, N., White, M., Patel, M., Chmielewski, J. and Walczak, K. (2003). AMS – Metadata for Cultural Exhibitions using Virtual Reality, In *Proc. Dublin Core Conference* (DC2003), Seattle, Washington, USA 2003.

Patel, M., Walczak, K., White, M., and Cellary, W. (2003). Digitisation to Presentation-Building Virtual Museum Exhibitions. *Proceedings of the International Conference on Vision, Video and Graphics*, Bath, UK, pp. 189-196.

Perrot, P. (1977). *The Smithsonian Experience*, Smithsonian Institute.

Schneiderman, B., and Plaisant, C. (2005). *Designing the User Interface: Strategies for Effective Human-Computer Interaction*-4th Edition, Addison-Wesley.

Sylaiou S., Almosawi A., Mania K., White M. (2004). Preliminary Evaluation of the Augmented Representation of Cultural Objects System. Proceedings of the *10th International Conference on Virtual Systems and Multimedia, Hybrid Realities-Digital Partners, Explorations in Art, Heritage, Science and the Human Factor, VSMM 2004 Conference*, 17-19 November 2004, Softopia Japan, Ogaki City, Japan, pp. 426-431.

Sylaiou S., Karoulis, A., and White M. (2006). Evaluation of the Augmented Representation of a Cultural Objects System. (submitted)

Tsapatori M. (2003). *ORION Research Roadmap, Evaluation and Assessment*, Object Rich Information Network (ORION), Deliverable 8, http://www.orion-net.org/Admin/LibraryLoc/file28.pdf (Retrieved 20 April 2005).

Wojciechowski, R., Walczak, K., White, M., and Cellary, W. (2004). Building Virtual and Augmented Reality Museum Exhibitions. *Proceedings of the Web3D 2004 Symposium - the 9th International Conference on 3D Web Technology, ACM SIGGRAPH*, Monterey, California (USA), pp. 135-144.