Proc. 5th Eurographics Ireland Chapter Workshop, Eurographics Ireland Workshop Series, Volume 3, Eds.: D. Murphy and J.O'Mullane, 9th Sept, Cork, Ireland, 16-23, (2004). ISSN: 1649-1807.

Using XML Technologies to Present Digital Content with Augmented Reality

N. Mourkoussis, F. Liarokapis, A. Basu, M. White, P.F. Lister

University Sussex, Department of Informatics,
Centre for VLSI and Computer Graphics, Falmer, BN1 9QT, UK

Abstract

This paper presents a low cost architecture that extensively uses XML technologies to present generic digital content using the web and augmented reality. We describe our client-server architecture with particular emphasis on unique XML schemas that are used to design a generic XML repository and illustrate its use with two different application scenarios. Our solution allows management through XML and publication to a visualisation client supporting both virtual and augmented reality integrated with standard web browsing. Two application scenarios have been developed to illustrate the effectiveness of the system.

Categories: Virtual and Augmented Reality, Computer Graphics, XML, Web Services.

1. Introduction

This paper presents our work and research findings on developing our low cost client-server architecture and its components for visualising digital content over the Internet or an intranet using XML, Web3D, Virtual and Augmented Reality (AR). Following our goals for multi-spaces visualisation and user interaction we have developed a prototype system called ARCOLite. This system can provide users with an interactive visualisation of digital content using web-based presentations and extending them into AR presentations.

In the past years only a few systems that combine VR, AR and Internet technologies have been developed. OpenTracker is an AR system based on XML tools used for development, configuration and documentation [RS03]. Another example is MARS, an authoring tool for creating and editing 3D hypermedia narratives based on XML, VR and AR technologies [GF03]. In addition, 2D applications, like internet applications can control the models displayed in the AR environment [Reg02]. ARITI is another system that allows a user-friendly operation interaction with the virtual representation in order to achieve difficult tasks remotely [OMK*00]. However, most of the developed architectures are designed based on dedicated software and hardware systems mostly without making in-depth use of all technologies.

The motivation of this work is to allow users to navigate through web-based presentations and extend them into immersive table-top AR presentations. The key feature of AR visualisation is that digital information can be augmented in the physical space such as a table-top environment. The AR interface allows the user to examine and manipulate the digital information (a 3D representation of a real world object) in the context of the physical table-top space. This provides an ideal low cost, novel and effective learning experience because, for example, the user can examine some physical objects on the table-top and within the same space see how these objects are related spatially. That is, how the objects are arranged in three-dimensional space. The digital information can also be animated thus giving the user a greater understanding of the objects, their use and meaning.

The remainder of this paper is organised as follows. In Section 2, an overview of the system is described based on the conceptual architecture of the system. Section 3 provides a deeper technical discussion of this XML repository. In Section 4, our XML Transformation engine is described which publish digital content to the visualisation client. In Section 5 the visualisation client is briefly analysed. Section 6 presents two different applications of this system; one used to present virtual museum exhibitions and another used to present an engineering education. Finally, section 7 concludes the paper and indicates future work.

2. System Overview

ARCOLite [WFM*04] has been originally derived from a more application-specific system called ARCO (Augmented Representation of Cultural Objects) [WMD*04]. ARCO is targeted at digitizing, storing, managing and publishing heritage digital consent to the Internet and AR using an Oracle database. However, ARCOLite is a more general and lightweight architecture for two reasons: it eliminates the database with an XML repository and it contains no non-standard technologies. ARCOLite offers methods for managing and presentation digital content in so-called virtual exhibitions (a combination of web, 3D and AR) based on technologies such as XML, Web3D, Java servlets, VRML97 [CB99] and the well known ARToolKit [KB99].

The requirements described in the introduction define the specification of the system architecture, which is illustrated in Figure 1. The architecture is divided in three conceptual parts: content management, exhibition server, and content visualisation.

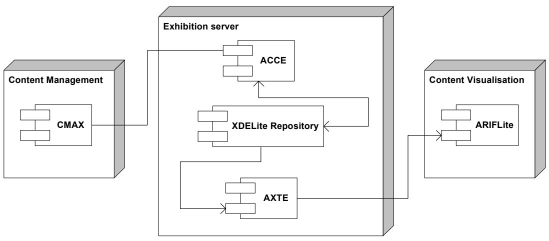

Figure 1: *Conceptual component diagram of ARCOLite*

The content management side consists of a web-based Content Management Application for XDELite (CMAX). CMAX is specified to provide a user-friendly interface for non-technical users to package raw multimedia data from local or network storage and metadata conforming to the XML data repository specification.

The exhibition server consists of three distinct components: the ARCOLite Content Creation Engine (ACCE), XDELite Repository and the ARCOLite XML Transformation Engine (AXTE).

The XDELite Repository is an XML-based data store and associated multimedia data, defined by a set of XML schemas, which is published according to XSL stylesheets (or templates) contained in the repository. ACCE is a Java servlet-based framework with which CMAX communicates to allow the content creator to add, edit or delete digital content in the XDELite Repository. ACCE contains low-level functionality that allows CMAX to package raw data and presentations according to the XDELite schemas (refer to section 4). AXTE is a servlet that serves digital content presentations (web, virtual reality and augmented reality) to the content visualisation clients.

The Content Visualisation component consists of a visualisation client also known as ARIFLite, which is a C++ application encapsulating a standard web browser and an AR table-top environment. This visualisation client combines the web-based form of presentation with either virtual reality or augmented reality presentations. The web presentations may also be visualised in any standards compliant XHTML web browser (i.e. Internet Explorer, Mozilla, etc.).

3. XDELite Repository

XDELite is the central data repository of the ARCOLite system. It is comprised of multimedia data (text and binary), metadata, an XML framework that associates multimedia data with metadata creating 'information resources', and XSL stylesheets defining the layout and the functionality of the digital content presentations.

3.1. The Data Model

The ARCOLite system is based on the data model, illustrated in Figure 2, which consists of two related entities [WFM*04]: the information resource and the media object.

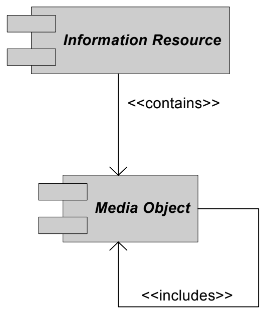

Figure 2: *Data model and metadata association*

We define an Information Resource (IR) as the representation of a physical artefact or an abstract concept stored in the ARCOLite system. An IR may be composed of one or more Media Objects (MO). The MOs are digital representations of the IR in a particular medium. Examples of MOs are 3D models, textual descriptions, static and panoramic (i.e. QuickTime VR) images [WMD*04]. Each MO type is identified with a distinct MIME type.

There are two main categories of MOs: simple and composite. Simple MOs correspond to media that can be represented in one data object—such as textual descriptions and static images that contain data directly. Composite MOs do not contain the data directly, but instead they are associated with a number of other MOs (either simple or composite) referred as children MOs. Composite MOs represent data objects with complex structure such as a VRML models, multi-resolution or panoramic images.

3.2. Metadata

A key element of the ARCOLite system is the specification of an appropriate metadata element set that underpins both the physical and technical aspects of the data utilised. We need to describe the IR, and preserve technical information of the processes involved for creating MOs. Accordingly, we have adopted the CO, MO and MOTs metadata schemas from a metadata element set collectively known as the ARCO Metadata Schema (AMS) [MWP*03, WMD*04]. The association of AMS descriptions to the ARCOLite data model entities is illustrated in Figure 3.

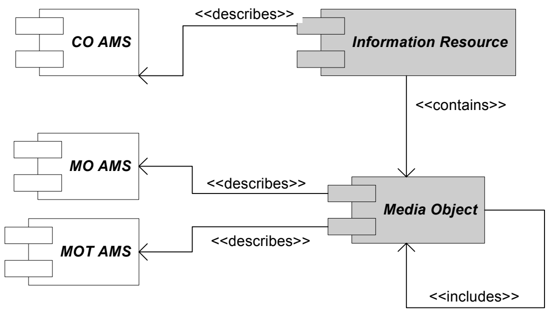

Figure 3: *Metadata association to data model*

Every information resource is described by the CO AMS schema that defines the physical characteristics of the information resource. The MO AMS portrays the technical characteristics of the MO whereas the MOT AMS schema provides supplementary information to the MO, in case it belongs to one of the predefined MIME types.

3.3. XML framework

The XML framework that associates the MOs and metadata consists of four distinct XML Schemas: 'IR', 'Presentation', 'Folder' and 'AR'.

The 'IR' aggregates data and metadata to create an information resource. The 'Presentation' is composed by a series of 'IR' instances, while a 'Folder' consists of several 'Presentation' instances and XSL stylesheets. Finally, the 'AR' schema defines which MOs will be transferred to ARIFLite in order to be used in the table-top AR environment. The 'AR' is a sub-set of the 'IR' schema, where its instance is created dynamically on the ARIFLite user request. Figure 3 illustrates the conceptual relationship of all the XML Schemas except of the AR one.

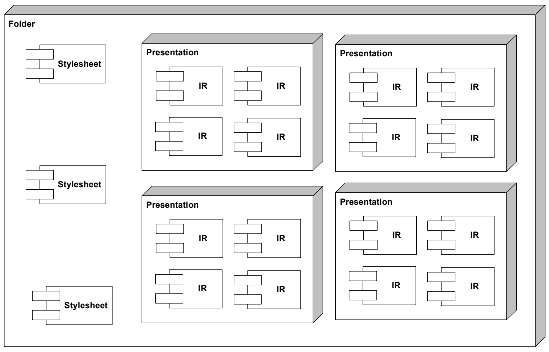

Figure 4: *XDELite schemas conceptual relationships*

Every IR must be comprised of an entry id and a data element. The entry id uniquely identifies the IR, whereas the data defines the thumbnail, the metadata schema, its instance and the MOs. Each MO must be comprised of a new entry id and data elements. Similar to IR, the entry id uniquely identifies the MO, whereas the data defines the media file, the metadata schema and its instance. The relationships of data that define an IR in ARCOLite are illustrated in Figure 5 below.

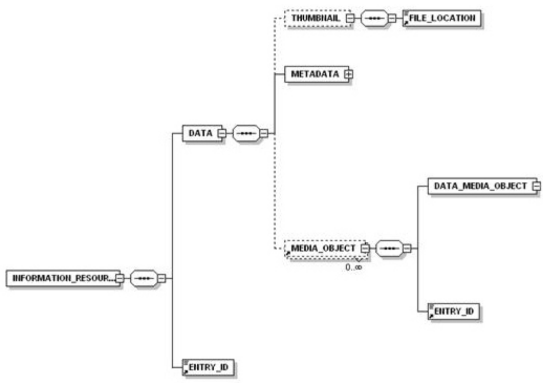

Figure 5: *IR schema data relationship*

In the same way, the other three schemas have been defined. More specific, Figure 6 presents the relationships of data that define a 'Presentation' schema in ARCOLite.

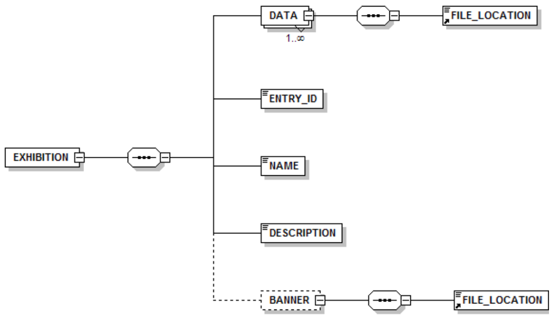

Figure 6: *Presentation schema data relationship*

In addition, the relationships of data that define a 'Folder' schema are presented in Figure 7.

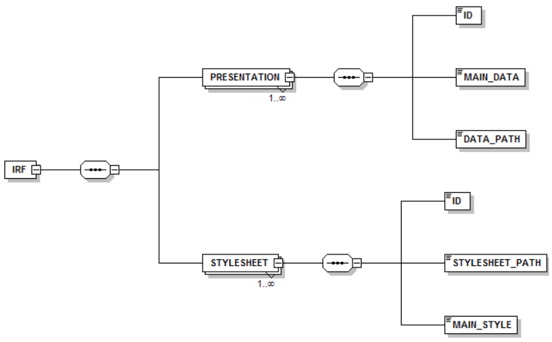

Figure 7: *Folder schema data relationship*

Finally, the relationships of data that define an 'AR' schema are illustrated in Figure 8.

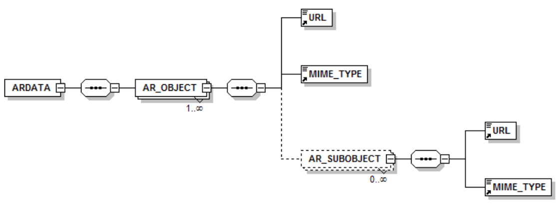

Figure 8: *Data relationship diagram in the AR Schema*

This XML framework (XDELite) functions in conjunction with the ARCOLite Transformation Engine (AXTE) to publish the data to the visualisation clients (ARIFLite).

4. ARCOLite Transformation Engine

AXTE is a Java servlet that provides the communication between the visualisation client and the XDELite Repository. Figure 9 illustrates the architecture of AXTE in the form of UML state-chart diagram representing its functionality. AXTE operates in four mutually exclusive modes known as 'MODE_XFROM', 'MODE_BINMED', 'MODE_BINTPL' and 'MODE_XFORM_PLAIN'.

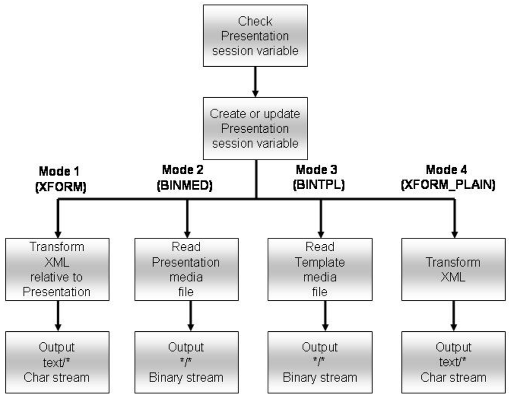

Figure 9: *AXTE state-chart*

The presentation templates written in XSL determine which mode is required in the course of processing the XML data or binary data. AXTE optionally prohibits the web browser from caching any output by adding cache expiration headers conforming to sections 13 and 14 in HTTP 1.1 RFC 2068. This is required to maintain consistency between multiple presentations served within the time limit of one browser session.

In 'MODE_XFROM', AXTE uses persistent session information to decide which presentation templates (XSL stylesheets) and 'Presentations' (XML instance files) are to be used during XSL processing. Using 'MODE_XFORM', AXTE delivers XHTML presentations rendered XHTML compliant visualisation clients.

'MODE_BINMED' serves binary stream of presentation multimedia content to the visualisation client. It uses persistent session information to decide the media file and its parent presentation before copying the contents of the file to the servlet output stream. On the other hand, 'MODE_BINTPL' delivers binary stream of multimedia media content relevant to the presentation templates (i.e. images, static script files and static cascading style sheets). It uses persistent session information to identify the presentation template directory.

AXTE uses 'MODE_XFORM_PLAIN' to deliver XML by transforming XML files (i.e. 'IR', 'Presentation' and 'Folder') through an XSL stylesheet. However, it does not use servlet sessions for persistent storage of path information for the XML files or the XSL stylesheets; instead all links to the data XML file and the XSL style sheet are treated relative to the servlet context path.

Thus with this combination of mutually exclusive modes, AXTE can act as a XSL transformation engine as well as a media content server.

5. AR Interface (ARIFLite)

ARIFLite is a visualisation client that allows communication between the web browser and an AR table-top environment. Based on techniques presented in [SM01] we have defined a way to dynamically attach digital information to the visualisation clients. In terms of operation, it requests digital content from the XDELite repository through AXTE and renders this data in both the web and the AR environment. Figure 10 illustrates a component diagram of ARIFLite.

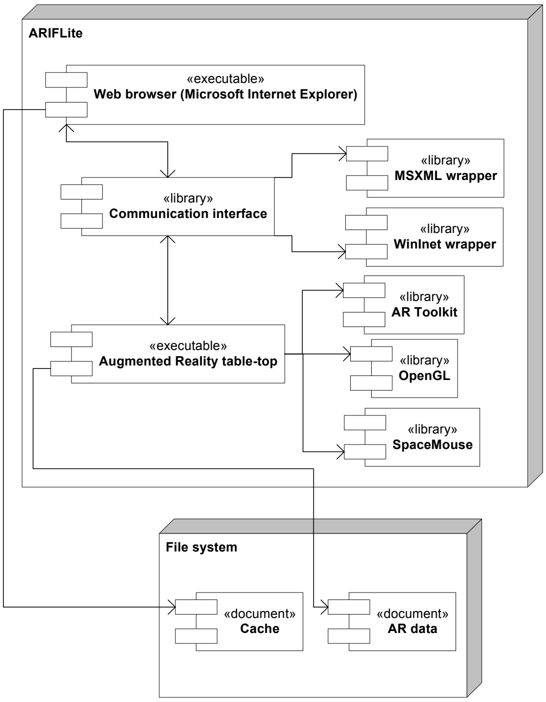

Figure 10: *Component diagram of ARIFLite*

ARIFLite consists of three major components; the web browser, the communication interface and the AR table-top interface. Each component is briefly presented in the following sections.

5.1. Web Browser

To make the digital information accessible to users all over the world, it needs to be distributed in an Internet-compliant way [CIG*01]. The web browser is an embedded Microsoft® Internet Explorer which communicates with AXTE (Figure 1) in order to visualise the XDELite repository. The retrieved digital content can be rendered with the aid of XSL stylesheets either on the web browser or in the AR table-top environment. By changing content and presentation style using different XSL stylesheets the visualisation can be altered in a number of different ways. For instance, Figure 11 illustrates how to make use of an archive lookup of a virtual cultural heritage exhibition presentation style that contains pictures and metadata.

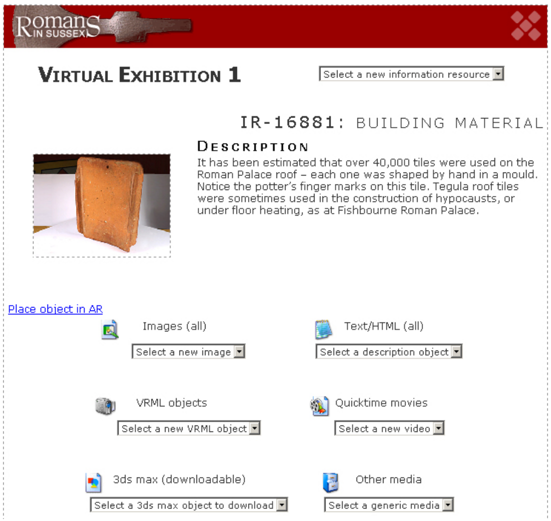

Figure 11: *Web-based heritage archive*

Looking at a more specific example (Figure 12), one can see for the web visualisation of an engineering component – a piston in this figure. The digital content contains, in this case, textual descriptions, thumbnail images and a 3D model but in the near future it will also include video animations and appropriate sounds.

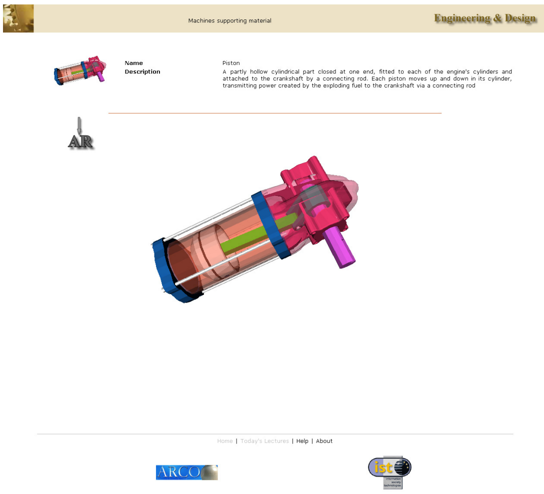

Figure 12: *Web visualisation of engineering content*

The 3D model (Figure 12) can be manipulated and viewed from all angles on the web page, which is a distinct advantage from looking at a simple image and reading a description. Furthermore, the biggest advantage of this technique is that different XSL stylesheets or templates can be used to present the same digital content in different ways for different audiences.

5.2. Communication Interface

The communication interface provides connectivity between the web browser and the AR components. Splitting the web browser and the AR interface can provide independent levels of control and implementation for three activities [SFT*99]: for rendering and displaying the digital information, for collaborative environments and for support of sophisticated input devices. Based on this approach, each time the AR interface is activated an 'AR' XML instance file is dynamically generated according to the 'AR' XML schema when the ARIFLite user clicks in a dedicated hyperlink (encapsulating a pseudo-protocol called ARIF) on the web browser, known as the "AR button".

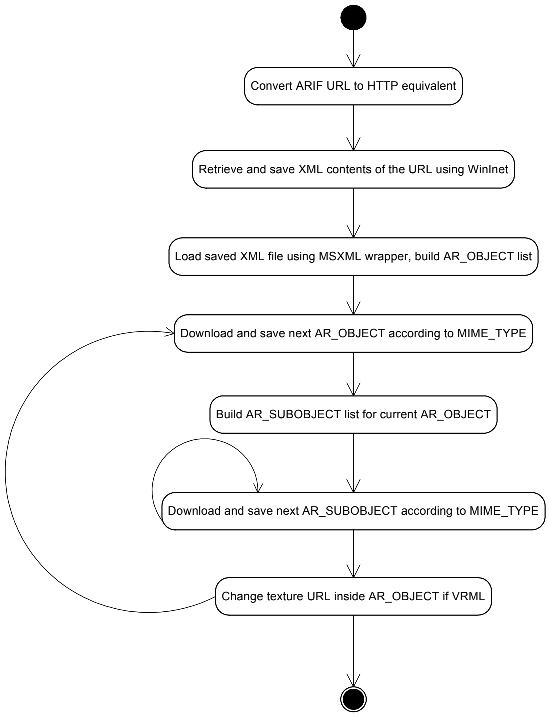

Figure 13: *UML state-chart diagram for the ARIFLite communication interface*

The ARIFLite communication interface is triggered by the ARIF pseudo-protocol but then it is handled by internal data structures. The functionality of this communication interface is best described in Figure 13. The pseudo-protocol URL is used to download a dynamically generated 'AR' XML instance file, which is parsed in several iterations by the underlying XML wrapper component. In this way, each and every media object and related sub-media object associated with the AR object (any media object) that is requested to be placed on the AR table-top interface are downloaded and stored.

5.3. AR table-top interface

The purpose of the table-top AR interface is to provide an interactive visualisation and also to provide a meaningful form of interaction with the digital information. This component of ARIFLite (Figure 10) consists of three software components wrapped around a C++/MFC framework [WFM*04] including: the ARToolKit 2.52, the Magellan SpaceMouse SDK and the OpenGL 1.2 API.

The ARToolKit tracking libraries were used for developing our AR interface. These libraries provide functionality for detecting the camera position and orientation and registering the digital information into predefined physical marker cards located in the real world (i.e. the physical table-top environment). In addition, ARToolKit is implemented in such a way so that it provides means of manipulating the digital information by physically moving physical marker cards in the real environment. Based on this, we have extended the techniques to perform a set of two different operations. First, we have carefully designed a set of marker cards to create a distinctive link between the user and our visualisation scenarios. Secondly, in a similar way to [LWL04] we have implemented means of manipulating the digital information in a realistic manner so that users do not feel unrelated with the environment.

For the visualisation a basic but robust computer graphics (CG) rendering engine based on C++ and OpenGL was implemented. The engine currently can support augmentation of 3D objects in VRML format [CB99] and textual information only, but in the future we plan to add sound and static pictures. The users can easily manipulate the augmented information based on a set of implemented transformations (i.e. rotations, translations and scaling). In addition, other important rendering techniques like interactive lighting and colouring may be used to manually 'correct' the visualisation to match reality.

Based on this software framework, users can intuitively explore digital representations in various ways. As in [BKP01], ARIFLite can handle multiple objects and assign it to multiple markers. In this way users can control the visualized information in a more efficient way because there is no need to return back to the web browser component in order to assign new objects.

To increase the level of immersive nature of the system the SpaceMouse SDK has been incorporated. This 3D interaction device offers users the ability to manipulate the digital information in six degrees-of-freedom (DOF). In addition, it includes a nine-button interface that was used to assign transformations (rotations, translations, and scaling) and other operations like clipping using an invisible plane, interactive lighting, etc. In this way, the standard keyboard may be eliminated or used as an alternative solution.

6. Applications of the ARCOLite system

This section illustrates how ARCOLite system can be used to support two different application scenarios for presenting digital content. The first is focused on presenting a museum exhibition while the second on teaching an engineering course.

6.1. Presenting museum exhibitions

For this scenario ARCOLite has been configured for a virtual museum application that allows users to interact with digital content using virtual and AR [WFM*04]. Traditional operation of museums may be enriched by having an online (3D) presence. Figure 14 presents in four screenshots the operation of the system configured for the virtual museum presentation.

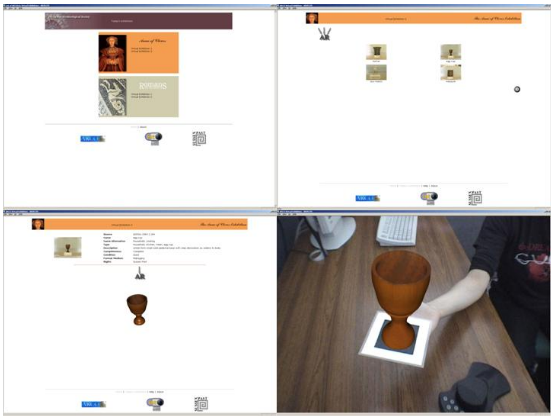

Figure 14: *Example virtual museum presentation*

Museum visitors can select a specific exhibition, which they are more interested (top-left window of Figure 14) and get thumbnail information about this category of artefacts (top-right window). Next, they can access more information about a specific artefact and interact with its VRML (virtual reality) representation using the mouse (bottom-left window). Finally, the visitors can move from web towards 3D visualisation and interaction by extending the presentation into our AR table-top environment (bottom-right window).

6.2. Teaching an engineering course

Based on the principles of the museum application scenario, ARCOLite can be easily reconfigured for an educational application that allows users to interact with 3D web content using virtual and AR [LMW*04]. A lecturer's traditional delivery can be enriched by viewing multimedia content locally or over the Internet, as well as in a table-top AR environment (Figure 15).

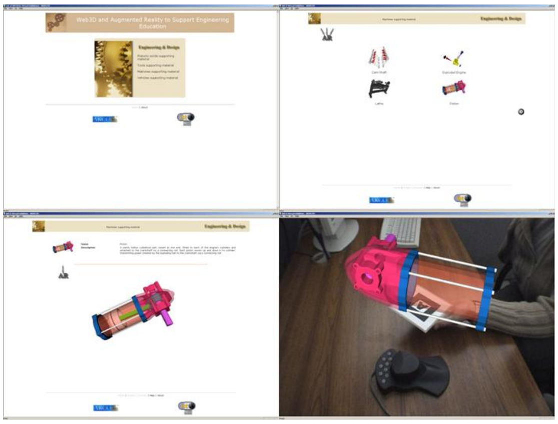

Figure 15: *Screenshot of the engineering education application*

In Figure 15, four web screenshots show the digital content relating to the example piston. More specific, the bottom-right image illustrates a 3D model of the piston has been presented in the AR view and the user is now examining the piston more closely. The user can now extract more information like its size relative to other physical objects. Such a combination of presenting digital content in web form and in the physical world offers the user a greater learning opportunity over other methods because it extends information from the 2D web pages into 3D content and then moves that 3D content into the users physical learning space.

7. Conclusion

ARCOLite provides an innovative system that allows users to interact with digital content using virtual and AR embedded in a Web3D environment. ARCOLite provides an architecture that can be configured and used in a diverse range of applications, by changing the theme of the digital content and the style of the presentation stylesheets stored within the XDELite repository. This enables us to explore the potential benefits of 'Web3D' and AR technologies in diverse application domains.

Some major improvements that could be implemented in the future includes the refinement of CMAX and ACCE (see Figure 1) and the addition of new XML schemas to define more interactive presentation styles such as virtual and augmented reality quizzes. Furthermore, we aim to implement new ways of interaction in the AR component using VR sensor devices such as the inertia cube.

Acknowledgements

This research was funded by the EU IST Framework V programme, Key Action III-Multimedia Content and Tools, Augmented Representation of Cultural Objects (ARCO) project IST-2000-28366. Special thanks to

Panagiotis Petridis and Joseph Darcy for providing the digital content and Maria Sifniotis for designing the visual layout of the stylesheets.

References

[BKP01] Billinghurst M., Kato H., Poupyrev I.: The MagicBook: A Transitional AR Interface, *Computer Graphics*, 25, 2001, 745-753.

[CB99] Carey R., Bell G.: The Annotated VRML 97 Reference. *Addison Wesley Longman, Inc.*, 1999.

[CIG*01] Cosmas J., Itegaki T., Green D., et al.: 3D MURALE: A Multimedia System for Archaeology. *ACM Siggraph Symposium on Virtual Reality, Archaeology and Cultural Heritage (VAST'01)*, Greece, 2001

[GF03] Güven S., Feiner S.: Authoring 3D Hypermedia for Wearable Augmented and Virtual Reality. *In Proc. ISWC '03 (Seventh International Symposium on Wearable Computers)*, White Plains, NY, October 2003, 118-226.

[KB99] Kato H., Billinghurst M.: Marker Tracking and HMD Calibration for a Video-based Augmented Reality Conferencing System. *In Proc. 2nd IEEE and ACM International Workshop on Augmented Reality*, San Francisco, USA 1999, 85-94.

[LMW*04] Liarokapis F., Mourkoussis N., White M., et al: Web3D and Augmented Reality to support Engineering Education. *In press, World Transactions on Engineering and Technology Education*, 2004.

[LWL04] Liarokapis F., White M., Lister P.F.: Augmented Reality Interface Toolkit, *In IEEE Proc. International Symposium on Augmented and Virtual Reality*, London (Jul. 2004), 761-767.

[MWP*03] Mourkoussis N., White M., Patel M., et al.: AMS – Metadata for Cultural Exhibitions using Virtual Reality, *In Proc. Dublin Core Conference (DC2003)*, Seattle, Washington, USA 2003.

[OMK*00] Otmane S., Mallem M., Kheddar A., et al.: Active virtual guides as an apparatus for augmented reality based telemanipulation system on the Internet. *In Proc. 33rd Annual Simulation Symposium*, 2000.

[Reg02] Regenbrecht H.T.: Interaction in a collaborative augmented reality environment. *In Proc. Conference on Human Factors in Computing Systems archive CHI '02 extended abstracts on Human factors in computing systems*, Minneapolis, Minnesota, USA, 2002.

[RS03] Reitmayr G., Schmalstieg D.: Location based applications for mobile augmented reality. *In Proc. Fourth Australian user interface conference on User interfaces*, Adelaide, Australia, Feb. 2003, 65-73.

[SFT*99] Strauss W., Fleischmann M., Thomsen M., et al.: Staging the space of mixed reality – Reconsidering the concept of a multi user environment. *In Proc. 4th Symposium on the virtual reality modeling language (VRML)*, 1999, 93-98.

[SM01] Sinclair P., Martinez K.: Adaptive Hypermedia in Augmented Reality. *In Proc. of the Third Workshop on Adaptive Hypertext and Hypermedia at the Twelfth ACM Conference on Hypertext and Hypermedia*, Denmark, August 2001, 217-219.

[WFM*04] White M., Liarokapis F., Mourkoussis N., et al.: ARCOlite-an XML based system for building and presenting Virtual Museums using Web3D and Augmented Reality. *In IEEE Proc. Theory and Practice of Computer Graphics 2004*, Bournemouth, (Jun. 2004), 94-101.

[WMD*04] White M., Mourkoussis N., Darcy J., et al.: ARCO – An Architecture for Digitization, Management and Presentation of Virtual Exhibitions. *In IEEE Proc. 22nd International Conference on Computer Graphics*, Crete, (Jun. 2004), 622-625.