IST-2002- 507382

EPOCH

Excellence in Processing Open Cultural Heritage

Network of Excellence

Information Society Technologies

D3.1.1: Overview of CH related IT research, related to stake-holder needs and the position of Europe therein

Due date of deliverable: 15 March 2005
Actual submission date: 29 April 2005

Start date of project: 15 March 2004
Duration: 4 Years

University of Leuven

Project co-funded by the European Commission within the Sixth Framework Programme (2002-2006)		
Dissemination Level		
PU	Public	X
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

Document Change Log

Each change or set of changes made to this document will result in an increment to the version number of the document. This change log records the process and identifies for each version number of the document the modification(s), which caused the version number to be incremented.

Change Log	**Version**	**Date**
Framework of this document	0.1	01/03/2005
First draft	1.1	25/03/2005
Final draft	2.1	25/04/2005

Table of Contents

DOCUMENT CHANGE LOG 3

TABLE OF CONTENTS 4

1. INTRODUCTION 6

2. WORK PACKAGE 3: JOINTLY EXECUTED RESEARCH 8

2.1 Timing and objectives 9

2.2. Activity 3.1: Coordinate research activities 9

3. OVERVIEW OF STAKEHOLDER NEEDS 11

3.1. Structure of cultural heritage needs 12

3.2. Most frequently recurring stakeholder needs 14

4. MULTILINGUAL AND SEMANTIC DATA PROCESSING 21

4.1. Systems, tools and technologies 21

4.2. Future directions and EPOCH's contribution 32

4.3. References 32

5. DATABASES AND KNOWLEDGE MANAGEMENT 34

5.1. Systems, tools and technologies 34

5.2. Future directions and EPOCH's contribution 42

5.3. References 45

6. MOBILE WEARABLE AND AMBIENT SYSTEMS 46

6.1. Systems, tools and technologies 46

6.2. Future directions and EPOCH's contribution 47

6.3. References 52

7. RECORDING AND DATA PRESENTATION 54

7.1. Systems, tools and technologies **54**

7.2. Future directions and EPOCH's contribution **79**

7.3. References **80**

8. VISUALISATION AND RENDERING **83**

8.1. Systems, tools and technologies **83**

8.2. Future directions and EPOCH's contribution **91**

8.3 References **92**

9. MULTI-MODAL INTERFACES **94**

9.1. Systems, tools and technologies **94**

9.2. Future directions and EPOCH's contribution **125**

9.3. References **127**

10. VIRTUAL HUMANS AND OTHER AVATARS **130**

10.1. Systems, tools and technologies **130**

10.2. Future directions and EPOCH's contribution **134**

10.3. References **139**

1. Introduction

"Europe's cultural and scientific knowledge resources are a unique public asset forming the colelctive and evolving memory of our divers societies and providing a solid basis for the development of our digital content industries in a sustainable knowledge society"

(Conclusions of the expert meeting of Member States, Lund, Sweden, April 2001)

European cultural heritage offers a wide range of possibilities for creating virtual environments, which can enable people to perceive different places, people, objects and events during different periods of history. All of these have embodied European local cultures and traditions through time and in turn they shape the developing culture of current societies. Therefore, their adequate communication and presentation can promote and provide a better understanding of the different European cultures.

In the last 10-15 years there has been a proliferation of organisations and projects applying IST solutions to cultural heritage applications. On the one hand this has grown from groups of technologists examining computationally interesting problems and requiring data for real cultural heritage artefacts and sites in order to demonstrate results. On the other hand groups of cultural heritage professionals have found computational tools which appear to allow new tools to explore their data and new opportunities to communicate results.

This situation extends into the arena of professional groupings, conferences and other organisations. There are a great number of organisations and professional groupings now involved in applying technological tools to assist in the archiving, preserving, and modelling of cultural heritage. VAST2001 (a symposium on Virtual Reality, Archaeology and Cultural Heritage) was sponsored by 4 EU projects in cultural heritage, and held in co-operation with 3 professional groupings active in related fields.

There are also dozens of universities which have carved out a focus in this domain. Finally, there are only a few companies that are focused primarily in this area – which is indicative of the scarcity of funding (Addison 2001).

This deliverable will focus on cultural heritage related IT research. It will clarify the position of Europe herein and relate it to the stakeholder needs.

2. Work package 3: Jointly Executed Research

This work package will guide and cross fertilise the research activities of the partners in the Network (and hopefully of others) in order to ensure maximal relevance for the Cultural Heritage domain, high quality, as well as cohesion and complementarity among these activities. A first activity therefore consists of Coordination of the Jointly Executed research (WP3.1).

A second activity aims at Establishing and Maintaining a Common Infrastructure (WP3.3) which will form a substrate upon which to build particular end-user applications and showcases. This activity will define the architecture, components and design guidelines for this common infrastructure. The goal is the creation of an integrated pipeline for producing applications involving digital versions of tangible cultural heritage.

A third activity (WP3.2) will develop New Tools to fill the gaps in this pipeline, or to create alternative technologies that are better suited for the Cultural Heritage domain. These activities will be based on the definition of the Research Agenda (WP2.5). The corresponding projects are henceforth referred to as NEWTON (= NEW TOols Needed) projects. The work package then implements the plan of action by soliciting and commissioning carefully targeted pieces of jointly executed research; and by monitoring and evaluating the components of the program in progress. Hence, the jointly executed research is expected to involve a mixture of integration of existing components and sub-systems and new tools to fill gaps in the integrated pipeline or improve functionality at specific bottlenecks.

The activities in this Workpackage are therefore:

- 3.1 Coordinate research activities
- 3.2 New tools
- 3.3 Common Infrastructure

2.1 Timing and objectives

The initial period of the JPA will be spent defining the priorities and picking the best candidate components for defining the common infrastructure. Part of this activity takes place with the reviews of available technologies and stakeholder needs in WP2. However part of it is done in WP3.1 and 3.3.

Under WP3.3 the initial period will be spent in experimentation with different promising technologies, which are candidates for the common infrastructure, to try and get a solid feel for the amount of additional effort that would be necessary to turn the current state of the technology into an integral part of a reliable infrastructure. The second important deliverable in the first year is the architectural design for the underpinning infrastructure, which will identify the required components and the interactions required between them. These interactions will then be compared with candidates for standardised interfaces, so as to allow development of plug-in components to the common infrastructure.

This gives rise to the following detailed JPA18 objectives for WP3:

- To define priorities and pick the best candidate components for the common infrastructure (in liaision with the reviews of available technologies and stake-holder needs in WP2). This will be done in Activities 3.1 and 3.3.
- To create an taxonomy of current projects and interests in the Network (Activity 3.1)
- To establish the principle of coordination and cooperation in research and build the confidence in relation to sharing and pooling ideas. (Activity 3.1)
- To define, process and commission the initial set of new tools projects (Activity 3.1/Activity 3.2)
- To begin work on the initial set of new tools projects (Activity 3.2)
- To define the architecture and functionality of the common infrastructure (Activity 3.3)

2.2. Activity 3.1: Coordinate research activities

Objectives

1. Organising regular consultation with Stakeholder Need (WP2.1) and Research Agenda (WP2.5) teams and giving feedback on technical feasibility and progress on the Common Infrastructure, in order to align technical and non-technical agendas.
2. Organisation of JER Concertation Meetings.
3. Creating links with other research projects – European and others – not funded through the Network, in a further effort to create critical mass. Similarly, acting proactively by assisting partners of the network in the bidding for research grants, at the European and national level

Description of work
1 Consultation with WP2 Based on Stakeholder Needs (WP2.1), WP2 will develop the Research Agenda (WP2.5). From WP3 additional input needs to be given, in order to balance desiderata against costs. A representative of WP3 will assist with these WP2 activities, in order to offer an integrated view on the results and status of ongoing JER (mainly the development of the Common Infrastructure during the first 18 months). ***2 Concertation Meetings*** Concertation meetings will be organised, in principle three during the first 18 months (in conjunction with the CAA and VAST conferences). The exact number during the 18 month period will depend on the time when these conferences are organised. ***3 Widening the network's scientific / technical impact*** The network will endeavour to maximally align research activities that go on outside its direct control. Related projects will be identified and contact will be established where possible. Information about relevant calls for project proposals will be collected and distributed among the partnership and via the website.

3. Overview of stakeholder needs

The goal of the EPOCH network is to develop and provide the proper ICT tools for the Cultural Heritage community and to improve the use and uptake of existing and new Cultural Heritage IT tools. The IT needs of all stakeholders within the Cultural Heritage community therefore first need to be defined and made public, so that the network can formulate a research agenda that benefits optimally the Cultural Heritage community, and that appropriate accompanying measures (training, brokerage, etc.) can be taken.

The first step to obtain such a common research agenda is to create an inventory of the needs of all stakeholders within the Cultural Heritage domain. As these needs are very diverse, it is important to create a detailed inventory of the needs of each stakeholder community—as identified by a consensus within that community itself. EPOCH's WP 2.1 tries to identify the most important stakeholder needs in the cultural heritage sector. A brief summary of their preliminary results is given below

3.1. Structure of cultural heritage needs

Phase 1: Data collection and processing
Textual and digital technology used for the acquisition of information about the artefact or site. This includes inventories, documentation, field recording, excavation data, monitoring, and post-excavation research and analysis.

Phase 2: Conservation / Preservation / Restoration
Technology used for on site or remote monitoring of the appropriate conditions of conservation and restoration. This includes preventive and active conservation, efficient inventorisation of protected or endangered cultural heritage resources, monitoring, security, urban and regional planning, development impact studies, and damage assessment.

Phase 3: Legal protection
Legal issues in relation with ICT. This includes protection of the collected and processed data, intellectual property and copyrights, technologies to prevent and fight the illicit traffic of cultural properties.

Phase 4: Management
This includes administrative functions such as planning, evaluation, reporting, and monitoring of site or museum use. It also includes improvement of ICT related services offered and the management of those services.

Phase 5: Training
Technology can be used to provide appropriate, continuous training and coaching for professionals and to introduce the visitors to ICT.

Phase 6: Interpretation / Education
Technology can play a major role in the presentation and the interpretation processes.

Phase 7: Valorisation
This topic covers the possible use of technology in the economic dimension of Cultural Heritage when it is considered as a resource for local development– as well as the enhancement of local, regional, national, or European identity.

<table>
<tr><td colspan="5"></td><td>1 Multi-lingual and semantic data processing</td><td>2 Databases and technology management</td><td>3 Mobile, wearable, ambient</td><td>4 Recording and data representation</td><td>5 Visualisation and rendering</td><td>6 Multi-modal interfaces</td><td>7 Virtual humans and other avatars</td></tr>
<tr><td rowspan="6">S3 Legal protec</td><td rowspan="6">S4 Mana</td><td rowspan="6">S5 Traini</td><td></td><td>A Project design and background research</td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td rowspan="3">S1 Data Collection and processing, S2 Conservation Preservation Restoration</td><td>B Data capture</td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>C Interpretation and analysis</td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>D Scolarly publication and archiving</td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>S6 Interpretation/E</td><td rowspan="2">E Public presentation</td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>S7 Valorisation</td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
</table>

This figure shows how these areas relate to the matrix used in WP 3.3.

3.2. Most frequently recurring stakeholder needs

The EPOCH stakeholder needs report proposed a first prioritisation of common needs among each step of the conservation and management process.

3.2.1. Data collection and processing

- Define policies, strategies, and standards for preservation, storage, documentation, exploitation and access of the heritage of digital information (sustainable system)
- Development of technology research programmes, either through external funding or incorporated within the routine tasks of the institution itself.
- Creation of a documented pipeline from source to presentation to improve data storage and data structuring.
- Digitisation of heritage resources and special archives or collections for planning purposes.
- Streamlined and organised data acquisition with clear objectives in the case of time constraints.
- Use of GIS for integration of heritage properties in overall urban planning process.
- Enhanced understandable access to relevant cultural heritage data bases and intelligent agents by members of associated communities, interested private heritage groups, visitors, and tourist professionals.
- Relationship between heritage education programmes and students performance in related discipline (history, science, art, etc.)
- Attention to safeguarding scientific accuracy, by making a clear distinction between fact and hypothesis.
- Data structures that contain multiple domain information, created by a multitude of researchers, physically present at different locations and multiple access using several channels for accessing CHIAs (interoperability across channels)
- Interface to link teachers, pupils, student and adult learners to a database containing appropriate and relevant information and to provide good practice examples to teachers and learning groups or individuals.

- Appropriate integration framework for exploiting all available information sources.

- Online access to cultural archives and intelligent gents to facilitate the collection of geographically dispersed information.

3.2.2. Conservation / Preservation / Restoration

- Information applications to link on-going conservation / preservation / restoration projects with legislation, internationally accepted conservation principles and practices, and international standards to monitor their success factors and evaluate existing policy guidelines. This should include conservation of both immovable heritage and other Cultural Heritage resources such as archives, libraries, manuscript collections, and art work.

- Creation of database of endangered sites, paths or artefacts within the local authority with links to other local public services (electricity, road building, traffic management, etc.) for efficient planning of local conservation, preservation, and restoration budget.

- Monitoring, anticipating, and planning restoration works will avoid long term destruction and reduce the risk of loss of authenticity or expensive emergency restoration (physical integrity and spatial setting) of a Cultural Heritage site.

- Public participation in the Environmental Impact Assessment through visualizations and analysis will enable the public to participate in the assessment of cultural heritage resources (Aarhus convent: considering of the use of electronic devices) and in the establishment in conservation / preservation / restoration priorities.

- Assistance in planning and policy decisions for more sustainable forms of community based cultural tourism development.

- Technology can provide educators with more detailed information about conservation procedures and standards, useful for educational programmes.

- Additional research to make the digital restoration and interpretation widely applicable.

- Support visitor management flows to ensure conservation of fragile or environmentally-sensitive cultural heritage sites.

- Access to all relevant legislation, charter, guidelines, to aid in the classification and protection of local heritage resources.

- Special respect for intangible heritage, which is often particularly sensitive and fragile.

- Access to information about restricted access on sites, changed classification of certain sites, environmental standards required (capacity of charge regulation), etc.

3.2.3. Legal protection

- Adoption of strategies and standards respecting internationally accepted conservation principles to ensure that Digitased Cultural Heritage can be preserved for the long-term.

- Policy on the use of ICT and practices that defines best practices, encourages the use of open source technology, and establishes common technological standards.

- Adequate attention to and provision of clear information on legal issues such as intellectual property and privacy.

- Need for a digital certificate of the originating institution for digitized documents.

- Access to official heritage policy formulation with the opportunity to stress the specific perspectives of associated communities not directly involved in local or national decision making processes.

- Means of access of the tourism stakeholder in the formulation of new heritage policies and standards to make their interests and perspectives known.

- Relevant for university level students in research is access to information on the wider legal and policy aspects of their work (in relation with the use of ICT). This information should also be made available in the formation of curricula of all cultural heritage courses.

- Adoption of standards regarding the clear indication/communication of where technology is delivering or enhancing an authentic “artefact” and where it is not.

- Access to all relevant legislation, charter, guidelines, to aid in the classification and protection of local digitised heritage resources.

- Provision of clear information on sources (copyright)

- Controlled access to digital archives, digital signatures and watermarking techniques, which can possibly be used together with other appropriate metadata descriptions that could be suitable for digital archives.

3.2.4. Management

- Need to understand the intellectual/ physical/ cultural/ geographical needs and expectations of audiences and to develop strategies to meet these needs as well appropriate communication/ education processes.

- Focus on interoperability, standards, and quality of services (i.e. usability, accessibility, monitoring of systems, maintenance, update of information, etc.)

- Close consultation, communication, and co-operation between site and museum managers, learners, teachers, cultural content providers, policymakers and technology provider is crucial from the very beginning of the project.

- Flexible and open data warehouses to facilitate exchange of data between specialist teams throughout the entire process of project development.

- On-line access to official policy documents and reports from the international heritage policy community.

- Internet applications to facilitate establishment of networks of local heritage authorities to share knowledge and tools; to exchange experience (success factors and failures); to promote results (studies, reports, guidelines); to avoid duplication of activities and waste of resources in simultaneous projects. Efforts should be merged, where possible, to maximise the impact on individual projects, creating a new and larger cultural community.

- Support back office functions such as human resources, finance and accounting, operations, ticketing, marketing / sales / E-commerce.

- Open access and ability to contribute to management systems by representatives of all relevant associated communities. Offer of informative and interactive IT services aimed at communication and participation of associated communities (forums, newsletters and Web bulletins)

- Visitors, expectations, experiences, reactions, cultural specificities, and physical needs should be assessed in detail, to create optimal cultural itineraries. Need of systems adapted to the user and so the need to cope with diversities in ability, culture and language. For example, the choice of technical devices for visitors should be assessed according to the physical preferences and capabilities of likely visitor group for example the reluctance or inability of some visitors to wear head mounted VR or audio devices for site interpretation should be considered in the selection of appropriate technology.

- Detailed monitoring of CH visitation statistics, economic outcomes and strategies better to assess the possibilities for success of planned heritage development / tourism initiatives as well as marketing and facilitation of package building.

- Use of IT tools and applications for visitor management flows (before and during the visit).

3.2.5. Training

- Professional training for members of the heritage policy community, the local heritage authorities, in data access and search, as well as digitalised glossaries of specific technical and legal terminology in all relevant languages.

- Appropriate lifelong learning processes: continuous guidance and on-site training programmes and courses with the objective of updating and informing of all levels of recent developments and innovations in the field for those creating, acquiring, using, or maintaining digital resources. This include Interested members of associated and host communities.

- Training of the operation front desk and guides.

- More efficient access to ICT by Cultural Heritage professionals for improving their competencies and related training and coaching. This can be accomplished in steps: recognition of existing resources; dissemination of information; exchange and promotion of collected data and experience.

- Access to information about employment opportunities in site management and education should be made available within local authorities.

- Need for relevant and practical information and pedagogical advice to support less experienced users.

- Need by Cultural Heritage institutions for new professionals in new fields with new skills on digital content creation, communication, and computer interfaces creation.

3.2.6. Interpretation / Education

- Democratic interactive open learning centre: implementation of compatible standards for Cultural Heritage and education (new learning) to accompany the digital content with lesson plans, concrete teaching tips and pedagogical guidelines, fitting into the national curriculum, by taking into account all aspects such as technology, content, and pedagogy.

- Need for considering interpretation and presentation from the very first stages of excavation or restoration onwards, using a wide range of sources (oral, written, research, local traditions), to create truly multidisciplinary study.

- Need to understand the policies and strategies (towards access) of the organisations/ institutions from which collections are digitized for an effective integration of ICT on-site.

- Need for long life cycle, adaptable and open interpretation systems.

- Anchoring interpretation technology within the relevant stakeholder groups and involved communities through direct communication (web magazines, forums, info) with local community, general public, general visitors, visitors with special interests, and experts, to disseminate Cultural Heritage, inform on new and existing heritage policies and stimulate public support for their implementation, understand perspective and understanding of associated communities and visitors

- Need to foster dialogue between cultures and to encourage and facilitate direct and active participation from visitors.

- Need to ensure inclusiveness of all types of visitors (visually and hearing impaired visitors, visitors with limited mobility and those with learning difficulties, etc.)

- Site interpretation must promote the enjoyment of the visitor as well as impart relevant information for understanding the site and its context (before, during and after the visit). This is particularly important for sites or museums used as educational resources.

- ICT should help to raise the visitors' awareness of heritage conservation issues before, during, and after the visit.

- Affordable and easy content management system as the kernel of interpretation systems (creation and integration of CMS – Content Management System)

3.2.7. Valorisation

- Data access to on-going public heritage projects to assess their compliance with official policies.

- Digital technology as a mean of soliciting and evaluating public response to propose heritage development plans: as an important component of policy decisions.

- Central internet portals and information exchange to promote the special character of regional heritage resources as a basis for promoting awareness and pride in distinctive local or regional identities.

- Avoidance of unrealistic or exaggerated assessments of the cultural value of sites or artefacts - to emphasise that the object or site and their history have a value independent of their presentation or marketed image.

- Need by European cultural institutions to use digitisation of cultural and scientific heritage as a means of preserving and valorising Europe's collective cultural and scientific patrimony; to safeguard sustainable and environmentally-conscious tourism; to support scientific research; and to contribute to the development of new digital content and service industries.

- In the planning process of heritage projects, groups and individual members of associated communities should have access to and input in the selection of sites and themes for valorisation in order to enhance the inclusiveness of heritage presentation and valorisation programmes.

- Need to better understand local context and issues to enhance more sustainable forms of community development.

- In line with government policy that museums and sites should be accessible resources to all, curators should look for ways to present information that helps the visitor to understand the subject and tourist professionals to develop effective and enjoyable tourism packages.

- Websites should enable educational opportunities such as allowing the visitor to search catalogues remotely and present exhibits not currently on display in the museum.

- Need to enhance multi-channel and multi-level presentations

4. Multilingual and semantic data processing

4.1. Systems, tools and technologies

In what follows is a review of the EPOCH relevant research areas, techniques and technologies involved in Natural Language Processing (NLP) for the design of Human-Computer Interactive (HCI) systems for Cultural Heritage (CH).

4.1.1. Speech Recognition

If spoken input is to be used as one modality, we must employ a Speech Recognizer (SR). Some aspects of what the SR delivers have a direct impact on the subsequent modules in the system. SR may be speaker dependent; in this case a training phase is required for the SR to get used to the speaker voice. Speaker independent SRs supporting wide varieties of languages are therefore preferable for Cultural Heritage applications. One key factor in choosing a SR is the Word Error Rate (WER), which is the percentage of words which are typically misrecognized by the SR, which is around 30% for state of the art SRs. Because CH applications are mostly domain-oriented, as opposed to task-oriented, it is also highly desirable that the SR allows the setting of recognition resource parameters (such as grammars) dynamically during run-time.

Typically, the format of the output of the SR is made up of words, out of vocabulary and pause symbols (OOV, PAUSE). It may also contain capital letters, common expressions combined as a single word ("Ducal_Palace"), short forms ("I'm"), hyphenated words ("short-sighted") and words between quotation mark (""). The SR produces a typical N-best list or Word Graph (WG) with confidence scores for words and for the whole sentence. Each path in a word graph represents an alternative sentence that the user could have said. Each word in a path has a confidence score, and each path is given an overall confidence score.

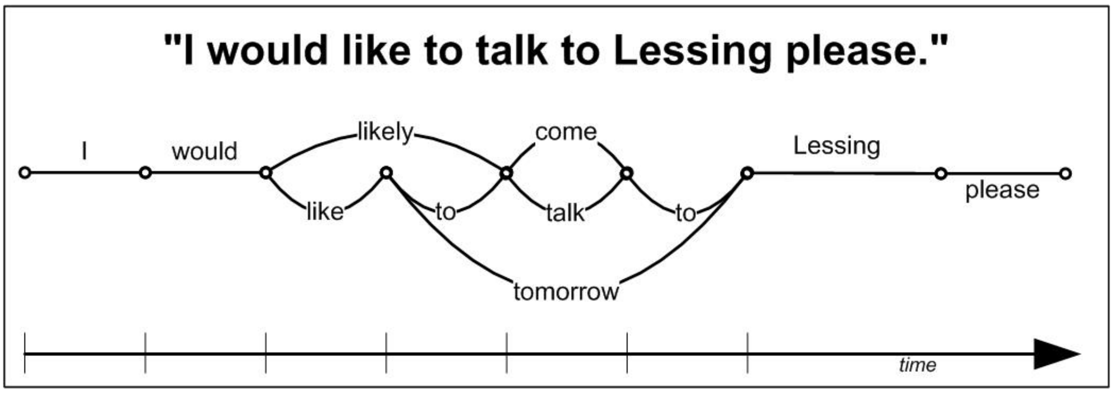

Figure 1: Word Graph for "I would like to talk to Lessing please"

Referring to Figure 1, the following five paths result from the WG:

1. I would like to talk to Lessing please.
2. I would like to come to Lessing please.
3. I would like tomorrow Lessing please.
4. I would likely come to Lessing please.
5. I would likely talk to Lessing please.

4.1.2. Semantic Parsing

The list of words coming out of the SR is to be interpreted semantically. The problem of finding a good interpretation of natural language utterances is particularly difficult to solve. *Semantic parsing*, as it is known, is described in (Allen 1995) as the process of mapping a natural language input (a sentence) to some structured meaning representation which is suitable for manipulation by a machine. In general, Natural Language Interfaces are difficult to build and must be tailored to each domain of application. Semantic parsing is challenging because it involves the concept of *Natural Language Understanding (NLU)*. Classical methods use hand-written rules and formal semantics to build up a suitable representation. Besides the heavy burden of finding appropriate rules for this task, formal semantics relies on the principle of *compositionality*, which does not take into account the non negligible idiomatic nature of natural language, especially in conversational speech. Not so long ago, corpus-based methods (Ng and Zelle 1997) received much attention for addressing these problems. They have been applied with success in areas like speech recognition (Rabiner 1989, Bahl et al. 1983), part-of-speech tagging (Charniak et al. 1993), text or discourse segmentation (Litman 1996) and syntactic parsing (Collins 1997, Manning and Carpenter 2000). They allow the construction of systems which satisfy desirable properties of NLP applications. These properties can be summarized as follow (Armstrong-Warwick 1993):

- Acquisition, i.e. automatically acquiring knowledge (domain specific or not) that would be necessary for the task
- Coverage, i.e. handling the potentially wide range of possibilities that could arise in the application
- Robustness, i.e. accommodating real data which may not be "perfect" (the noise factor) and still being able to perform reasonably well
- Portability, i.e. easily applicable to a different task in a new domain

Corpus-based methods rely on a training corpus which is, in general, a collection of sentence-meaning pairs (for the task of semantic parsing). Two main areas of corpus-based approaches for semantic parsing can be distinguished: the *Machine Learning* approach and the *Statistical* approach. In the machine learning approach, some learning algorithm is used to learn how to go from utterances to meanings. Statistical approaches collect a number of parameters from the training corpus to help a semantic parser discriminate between good and not so good meanings. Some approaches use both. Traditionally, methods used to construct semantic parsers very often involve the creation of rules by a knowledge expert. These hand-crafted rules make the parser incomplete, even for a specific domain. As knowledge to encode grew in size, handcrafted work became more and more difficult as we approach the so-called *knowledge engineering bottleneck*. The result was a very inefficient and fragile parser. More recent approaches tend to avoid this knowledge-engineering perspective in favour of an empirical, corpus-based approach where parsers are constructed through learning.

4.1.3. Dialogue Management (DM)

With the advent of domain-oriented Spoken Dialogue Systems (SDSs), the task of managing a conversation will probably undergo fundamental change in the next few years. It is one thing to convey a proper conversation in a task-oriented system, e.g. to help a customer to buy plane tickets or get information on current events, it is quite another to enter in a conversation about the effects of global warming or life in general. By constraining the input, task-oriented systems have achieved a satisfactorily level of functionality. Domain-oriented systems are in their infancy. Both types of conversations may be valuable for cultural heritage applications. A request about items being on display in a museum is a simple task, while discussing with a virtual historical character about its work is a domain-oriented dialogue. One important lesson for CH applications can be drawn from work by (Core et al. 2003) with regards to initiative in tutorial dialogues, where a proposed set of speech acts helps tracking down who has the *initiative*. In (Whittaker and Stenton 1988), *initiative* is defined as the control taking of the dialogue by one participant. More to the point, (Chu-Carroll and Brown 1998) differentiate dialogue initiative from task initiative. Dialogue initiative "tracks the lead in determining the current discourse focus" (p.6) and task initiative "tracks the lead in the development of the agent's plan" (p.6). Roughly speaking, besides the purely semantic content of dialogues in conversations, an important aspect of success in speech, measured by the user's satisfaction with the quality of the output, variation, humour, breath and depth of topics, comes from knowing who has the initiative and who should now have it. The main reason appears clear by looking at findings from (Core et al. 2003) which studied initiative management in two dialogue strategies: *didactic* tutoring and *socratic* tutoring. Socratic tutoring "is characterized by the use of questions and other hints to draw out answers from students having difficulty" (p.1), while in didactic tutoring, "the tutor points out the student's error and explains how to derive the correct answer" (p.2). What these findings show is that if we want the student to achieve as much learning as possible in a tutorial dialogue, then it is wiser to adopt a socratic way of carrying out the dialogue by asking more questions, with roughly one out of ten turns being led by the student. As a matter of fact, this will increase student verbosity, and make the dialogue more interactive. Finally, (Core et al. 2003) notes a positive correlation between interactivity and learning. In spontaneous conversation drawn from a large domain, the tutor is replaced by a machine and the goal of the machine dialogue management unit is more blurred (entertainment, education, information, etc.) where learning *per se* is not the final goal, but part of it. Because *verbosity*, *interactivity* and *learning* all contribute to make a dialogue successful, one can adopt a dialogue management policy inspired by (Core et al. 2003)'s findings for tutorial dialogues.

4.1.4. Speech Generation

Most work in Natural Language Generation has focused on written language, as opposed to spoken output (Biber et al. 1999). The study of speech generation has yet to achieve the same level of maturity of its close relative, and most of the spoken systems have tailored output originally intended for document creation, in other words written language, to be passed along to a plain text-to-speech synthesizer. Tailoring involved mainly that the text to be synthesized is augmented with basic sentence level prosodic information. However, there is more to this than meets the eye; to convey properly the intended semantics as well as the emotion of the speaker and its particular style, we need more than the bare words; *concept-to-speech* systems, as they have been termed, attempt to act directly on the final phonetic representation based on contextual, linguistic and semantic information to convey semantic, affective and stylistic information. Variable output can be achieved via stochastic generation.

4.1.5. Stochastic Generation

To produce human-like spoken output, one must take into account variability, and this is where language modelling comes into play. Language modelling using trigrams has been around for some time (Bahl et al. 1983) as well as tools for creating the models (Clarkson and Rosenfeld 1997). In recognition, language models are used to extract the most probable sequence of words given the acoustic evidence. In generation, the "evidence" we have is the utterance's class, a combination of topic and speech act. We are not looking for the most likely sequence, but rather for the generation of outputs according to the distribution of each gram in the corpus for that particular class. Pioneering work in this area is (Rudnicky and Oh 2000). The general idea of the generation algorithm is that words are generated randomly until the end of the sentence marker "</utt>" is generated. It should be clear that whenever a higher order n-gram model is used in generation, the fertility decreases while precision increases; using a lower order model has the exact opposite effect on precision and fertility. The arrows in Figure 2 illustrate the effect of lowering or raising the model's order on precision and fertility.

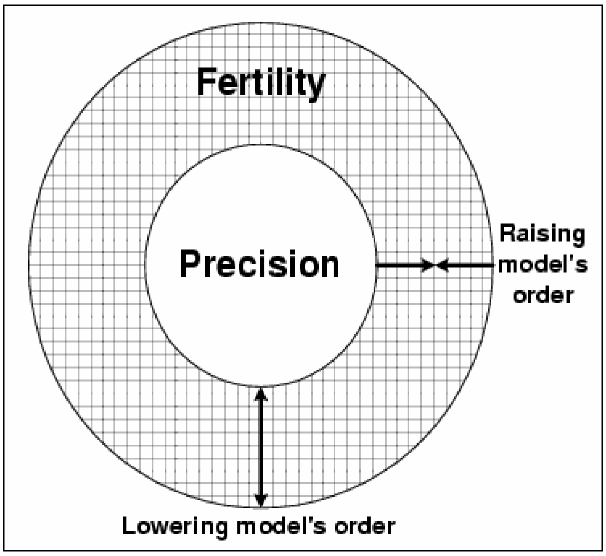

Figure 2: Precision versus Fertility

4.1.6. Multilinguality

Globalization means that interaction will be carried out in different languages.
The main problem with applications that interact in more than one languages is that *interpretation* must be preserved, not (literal) *meaning*. Consider the following German example, uttered by a virtual guide to seek confirmation from a user navigating and interacting through a virtual replica of a historical town: “Geht es bei Dir?”. Depending on the target language, the literal meaning rendered in other languages is more or less close to the intended interpretation. In French, automatic translation produced by a machine could give something like « *Va-t-il avec toi* ? » whilst *“Does it go with you?”* could be a possible English translation. French users would probably infer quickly the right interpretation “Ça va?”, but the English speakers would probably struggle to interpret the guide’s reply in the right way “*Is that alright with you?*”. The difficulty is that interpretation depends on context, but the good news is that interactive systems are considering more and more modalities, as we will see in the next section, which provide a richer source of contexts, for input interpretation and for output generation.

4.1.7. Multimodality

Modern generation systems for speech are set in a *multimodal* environment. In this type of environment, speech is not the only mean of interaction; point and click operations, for example, can be used along with speech. Because each modality interacts with, complements and to some extent modifies one another, modern speech generation must be studied in a multimodal setting. A long-term objective for dialogue systems is “sensory realism”, so that humans interacting with a machine could use their senses in a natural way. On one hand, visual and auditory are two modes of feedback in HCI that have been used extensively. On the input side, gesture and speech recognition have progressed significantly over the last few years. However, the touch, taste and smell modalities are yet to be used broadly in HCI. On the other hand, users do not employ modalities in the same way: for example, in the domain of searching newspaper texts, experiments (Klein et al. 2001) show that non-expert users benefit from (time wise) and prefer to combine speech and click operations over written and click operations.

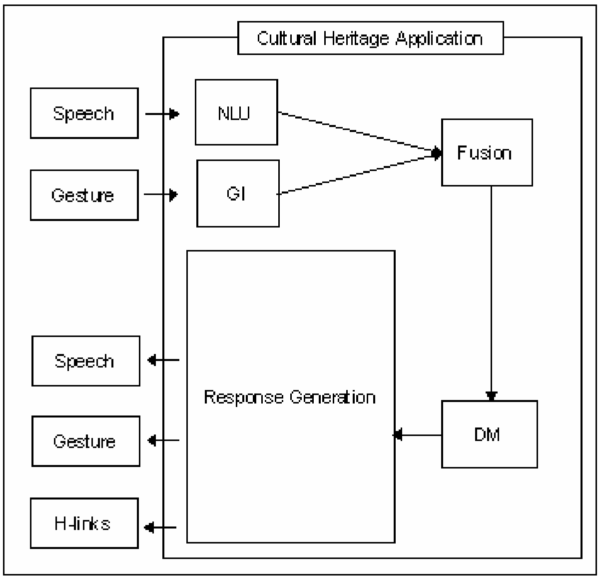

Figure 3: Integrating Modalities (GI = Gesture Interpretation)

Therefore, a major challenge for multimodal applications is the integration of those modalities. To be interpreted correctly and efficiently, parallel modes of expressions need to be integrated in a general and principled framework so that synchronization and further development fall in naturally in the unified model. A typical integration scheme for CH application may look like Figure 3.

4.1.8. Existing relevant projects for Spoken Dialogue Systems in Cultural Heritage

Three applications can provide a very useful framework for the development of Dialogue systems in the Cultural Heritage domain. (WYSIWYM) aims to allow domain experts to encode their knowledge directly. The (NECA) system, originally designed for combining situation-based generation of natural language and speech, gesture and emotions situated in social communication, provides a general model for the integration modalities and emotion during the integration. The (M-PIRO) project provides a direct example of how a system that generates descriptions of museum objects can be tailored to the user in terms of level of expertise, modality expressed, wording, phrasal complexity and language.

4.1.9. Putting it all together

In Figure 4, we show the data flow for a typical spoken interaction between a user saying “Do you like Harry Potter?” to a virtual tour guide, for which the guide strategy is simply to refocus the attention to a topic he is more familiar with: “Actually, I like von Goethe”.

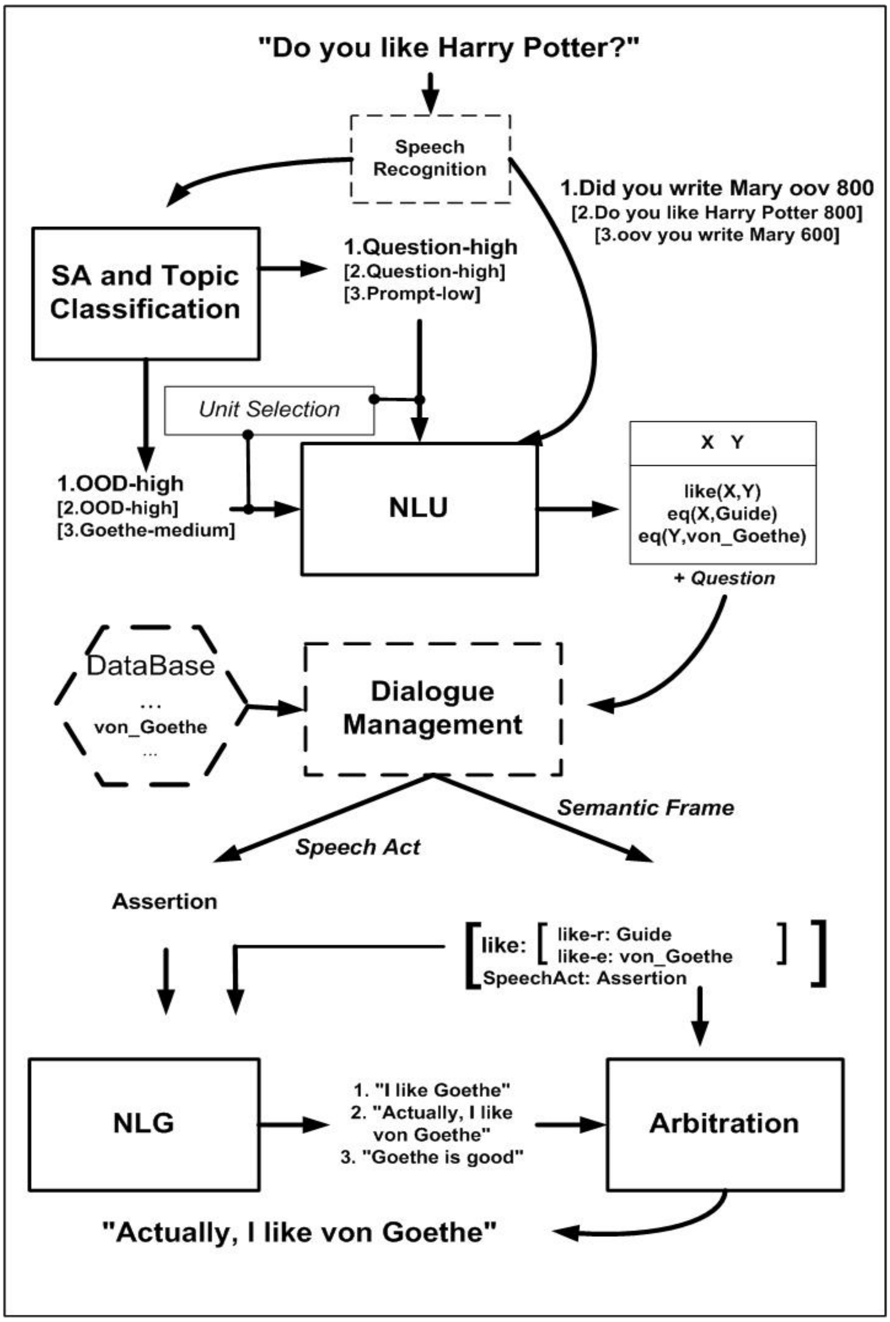

Figure 4: Data flow for a typical spoken interaction

4.1.10. State of the Art: Natural Language Generation in guided tour systems

AT&T NJFun[1]

“NJFUN is a spoken dialogue agent that allows users to access information about things to do in New Jersey, via a telephone conversation”. It is not multilingual and does not involve any Virtual Reality (VR). “NJFUN has served as a test bed for demonstrating the use of Reinforcement Learning for optimizing spoken dialogue management. It uses a speech recognizer with stochastic language models trained from example user utterances, and a TTS system based on concatenative diphone synthesis. Its database is populated from a public webpage to contain information about activities”.

IRST AL Fresco[2]

The IRST AL Fresco system is a museum guide that sits on a hand-held device. It takes visitors round an Italian museum, and is multilingual, but with no VR. “ALFRESCO is a prototype with the purpose to exploit the potentiality of integrating natural language with other communicative modalities”, so it is multimodal. “The approach relies on the technological possibility of moving in a rich information space with rapid interleaving of different information and different media, and on the high-level integration into a coherent information seeking dialogue”.

ILEX[3]

The ILEX system generates dynamic labels for visitors to a Scottish museum. It is a web based system, is not multilingual and does not involve any VR. The focus of the project is automatic text generation, in order to “produce descriptive, explanatory, or argumentative texts to accomplish various different communicative tasks”. The very dynamic nature of the labels “has a number of advantages, such as taking into account the visitor's level of expertise about the objects, as well as the discourse history---the objects which the visitor has already seen---so that information the visitor has already assimilated can be taken into account description of the object currently being viewed can make use of comparisons and contrasts to previously-viewed objects, while omitting any background information that the visitor has already been told”.

[1] Satinder Singh, Diane Litman, Michael Kearns and Marilyn Walker. Optimizing Dialogue Management with Reinforcement Learning: Experiments with the NJFun System. Journal of Artificial Intelligence Research(JAIR), 2002.

[2] http://tcc.itc.it/history/projects/alfresco/alfresco.html

[3] http://www.hcrc.ed.ac.uk/ilex/

M-PIRO[4]

M-PIRO is an extension of the ILEX project. While “ILEX served up personalized information objects”, M-PIRO main advance is that “it develops an authoring tool (developed at NCSR “Demokritos”) to help museum creators create and edit the domain-specific knowledge base and linguistic resources”. It is multilingual, uses a museum guide, but no VR, although one partner (“Foundation of Hellenic World”) “have been working on a new M-PIRO prototype that will embed the project’s technology in their immersive VR system”.

ICT[5]

On the other side of the Atlantic, one leader in research on speaking avatars is ICT in Marina del Rey, California. A lot of their applications are military and most (if not all) are unilingual, but there is a system in development called “Integrating Architecture” that “will provide an integrated infrastructure for fundamental research in the disparate areas of artificial intelligence (AI), graphics, sound, animation, and immersive display technologies, such as FlatWorld (the Mixed Reality Simulation Space) and the Virtual Reality Theater”, which could be relevant for EPOCH.

NICE[6]

The NICE project has just completed and a working system can be used by visitors in the Hans Christian Museum in Odense Denmark. “NICE aims to demonstrate universal natural interactive access, in particular for children and adolescents, by developing natural, fun and experientially rich communication between humans and embodied historical and literary characters. The communication consists of domain-oriented spoken conversation combined with 2D input gesture into a 3D dynamic graphics virtual world inhabited by the fairy-tale author Hans Christian Andersen and animated characters from his fairy-tale universe. For the first time, professional computer games technologies are joined with advanced spoken interaction, and speech recognition technology is specially developed for recognising the speech and spoken linguistic behaviour of children and adolescents”.

NECA[7]

“NECA promotes the concept of multi-modal communication with animated synthetic personalities. A particular focus in the project lies on communication between animated characters that exhibit credible personality traits and affective behaviour. The key challenge of the project is the fruitful combination of different research strands including situation-

[4] http://www.ltg.ed.ac.uk/mpiro/

[5] http://www.ict.usc.edu/

[6] http://www.niceproject.com/

[7] http://www.oefai.at/NECA/project/project.html

based generation of natural language and speech, semiotics of non-verbal expression in situated social communication, and the modelling of emotions and personality.” The i-Guide showcase was heavily based on the NECA architecture.

Text-to-Speech systems (TTS)

Name	DESCRIPTION
MBROLA	A free Phoneme-to-Speech system which includes many voices, among them English, French and German, for which a free Text-to-Phoneme is also available http://tcts.fpms.ac.be/synthesis/mbrola/
FreeTTS	A free TTS. “FreeTTS is a speech synthesis system written entirely in the JavaTM programming language. It is based upon Flite: a small run-time speech synthesis engine developed at Carnegie Mellon University. Flite is derived from the Festival Speech Synthesis System from the University of Edinburgh and the FestVox project from Carnegie Mellon University”. http://freetts.sourceforge.net/
Festival	A free TTS. “Festival offers a general framework for building speech synthesis systems as well as including examples of various modules. As a whole it offers full text to speech through a number APIs: from shell level, though a Scheme command interpreter, as a C++ library, from Java, and an Emacs interface. Festival is multi-lingual (currently English (British and American), and Spanish) though English is the most advanced”. http://www.cstr.ed.ac.uk/projects/festival/
Flite	A free TTS. “Flite (festival-lite) is a small, fast run-time synthesis engine developed at CMU and primarily designed for small embedded machines and/or large servers. Flite is designed as an alternative synthesis engine to Festival for voices built using the FestVox suite of voice building tools”. http://www.speech.cs.cmu.edu/flite/
ViaVoice	A multilingual commercial TTS. http://www.scansoft.com/viavoice/
Nuance	A multilingual commercial TTS. http://www.nuance.com/

Markups (Pirker and Krenn, 2002)

Name	DESCRIPTION
VHML	“A markup language for the representation of different aspects of avatars, such as speech production, facial and body animation, emotional representation, dialogue management, and hyper and multimedia information”. http://www.vhml.org
MPML	“MPML (Multimodal Presentation Markup Language) is an XML-based markup language developed to enable the description of multimodal presentation on the WWW based on animated characters”. http://www.miv.t.u-tokyo.ac.jp/MPML/en/
SSML	“The essential role of the markup language is to provide authors of synthesizable content a standard way to control aspects of speech such as pronunciation, volume, pitch, rate, etc. across different synthesis-capable platforms”. http://www.w3.org/TR/speech-synthesis/
SAPI TTS	The Microsoft TTS markup. http://www.microsoft.com
SML	The sub-part of VHML concerned with the markup for speech synthesis.
RRL	“A Rich Representation Language for the description of agent behaviour in the NECA project”. http://www.oefai.at/NECA/RRL/RRL_docs/RRL-FINAL.pdf

4.2. Future directions and EPOCH's contribution

Inside EPOCH it is proposed to strengthen the links with area 2, in order to apply CIDOC-CRM metadata for the creation of multi-lingual query capability. Semantic data processing should allow converting work from existing ontologies to CIDOC-CRM.

It is also proposed to strengthen the links with area 5 for developing modelling and authoring tools to create intelligent, knowledge-aware environments. A stronger link with area 7 could lead to better control of avatar timing and gesture.

Another aim is to generate better tailored responses by incorporating the context of the visitor.

To comply with these aims, additional support will be sought under call 5 for story-telling, emotion…

4.3. References

Allen, J. 1995. *Natural Language Understanding*. Benjamin Cummings.

Armstrong-Warwick, S. 1993. Preface to the Special Issue on Using Large Corpora. *Computational Linguistics*, 19(1):iii-iv.

Bahl, L., Jelinek, F., and Mercer, R. 1983. A Maximum Likelihood Approach to Continuous Speech Recognition. *IEEE Transactions on Pattern Analysis and Machine Int*., pages 179-190.

Biber, D., Johansson, S., Leech, G., Conrad, S., and Finegan, E. 1999. *Longman Grammar of Spoken and Written ENGLISH*. Pearson Education Limited, Edinburg, UK.

Charniak, E., Hendrickson, C., Jacobson, C., and Perkowitz, M. 1993. Equations for Part-of-Speech Tagging. *Proceedings of the Eleventh National Conference on Artificial Intelligence*, pages 784-789.

Chu-Carroll, J. and Brown, M. K. 1998. An Evidential Model for Tracking Initiative in Collaborative Dialogue Interactions. *User Modeling and User-Adapted Interaction*, 8:215-253.

Clarkson, P. and Rosenfeld, R. 1997. Statistical Language Modeling using the CMU-Cambridge Toolkit. *Proc. Eurospeech 1997*, pages 2707-2710.

Collins, M. 1997. Three Generative, Lexicalized Models for Statistical Parsing. In Cohen, P. R. and Wahlster, W., editors, *Proceedings of the Thirty-Fifth Annual Meeting of the Association for Computational Linguistics and Eighth Conference of the European Chapter of the Association for Computational Linguistics*, pages 16-23, Somerset, New Jersey. Association for Computational Linguistics.

Core, M. G., Moore, J. D., and Zinn, C. 2003. The Role of Initiative in Tutorial Dialogue. *10th Conference of the European Chapter of the Association for Computational Linguistics*.

Klein, A., Schwank, I., Généreux, M., and Trost, H. 2001. Evaluating Multi-modal Input Modes in a Wizard-of-Oz Study for the Domain of Web Search. Ann Blandford, Jean Vanderdonckt and Phil Gray eds, *People and Conputer XV - Interaction without Frontiers: Joint Proceedings of HCI 2001 and IHM 2001*, pages 475-483.

Litman, D. 1996. Cue Phrase Classification using Machine Learning. *Journal of Artificial Intelligence Research*, 5:53-95.

Manning, C. and Carpenter, B. 2000. Probabilistic Parsing Using Left Corner Language Models. In Bunt, H. and Nijholt, A., editors, *Advances in Probabilistic and Other Parsing Technologies*, pages 105-124. Kluwer Academic Publishers.

M-PIRO: http://www.ltg.ed.ac.uk/mpiro/

NECA: http://www.oefai.at/NECA/

Ng, H. T. and Zelle, J. 1997. *Corpus-Based Approaches to Semantic Interpretation in Natural Language Processing*. Special Issue on Empirical Natural Language Processing{AI Magazine Winter 1997, 184:45{64.

Hannes Pirker and Brigitte Krenn 2002, *Assessment of Markup Languages for Avatars, multimedia and multimodal system*. Neca Deliverable D9c, ÖFAI, http://www.oefai.at/NECA/publications/publication_docs/d9c.pdf

Rabiner, L. 1989. A Tutorial on hidden Markov Models and Selected Applications in Speech Recognition. *Proceedings of the IEEE*, 772:257-286.

Whittaker, S. and Stenton, P. 1988. Cues and Control in Expert-Client Dialogues. *Proc. of the 26th Annual Meeting of the Association for Computational Linguistics*, pages 123-130.

WYSIWYM: http://www.itri.bton.ac.uk/projects/WYSIWYM/wysiwym.html

5. Databases and knowledge management

5.1. Systems, tools and technologies

Lock (2003) provides a good introduction to the use of databases in cultural heritage. The earliest databases were ‘flat file’ and reflected the simple hierarchical data structure of a card index in which each record consists of a series of logically related fields, and each file consists of a series of logically related records. Although flat-file databases still find many applications within cultural heritage this is an inefficient data structure, not least because fields that are left empty still have to be stored and analysed, and repeating values in a field have to be stored and entered many times. Several alternative data models have been developed to improve efficiency, but the relational model is by far the most popular for structured data. Other solutions have been developed for databases that consist of large amounts of free text.

A relational database consists of a number of tables made up of rows and columns. Tables are linked by a primary key which is duplicated in tables to be joined. Relationships can be either one-to-one, or one-to-many. The order of rows is not significant, whereas the order of columns is significant (Smith 1991).

Relational databases offer a powerful set of data manipulation operators which must include the basic three: SELECT allows the production of a subset of rows according to specified criteria; PROJECT produces a specified subset of columns from a relation; JOIN allows the concatenation of two or more tables according to specified criteria and the above two operations on the resulting new relation. These plus other standard operators permit the construction of queries of infinite complexity.

There are two main approaches to communicating with the database. The first uses the concept of ‘forms’ which can be built from one or more relations to reflect any aspect of the data structure and be displayed on the screen for data input or retrieval. The second approach is to use a query language of which SQL (Structured Query Language) has become the international standard.

Both form-based and SQL queries can be saved for subsequent modification and reuse. The results of queries can be produces within various report templates so that catalogues and other documents can be easily produced. Most modern relational databases incorporate graphical output which will produce histograms, bar charts and other descriptive statistics as the result of a query. An exciting development is the ability to include raster images as fields for each record; this holds tremendous potential for cultural heritage even though at the moment images can only be looked at and not used analytically (see chapter X).

Flat-file and relational databases are designed to be efficient for data that are rigorously structured into tables with repeating records and repeating fields, such data may be coded. While these usually allow free text within specified fields they are not suitable for large volumes of textual data. Text database software has developed which now allows the querying of text as quickly and efficiently as traditionally structured data. The following are the two most common types.

Free-text databases are designed to work with completely unstructured large volumes of text. Instead of searching on the contents of fields, a free-text database will construct an index of every word which is a potential search criterion. Text to be included in such a database is likely to have been produced by word-processing which is then read into the database. During this process each word is checked against a ‘stop list’ of words not to be indexed. And, if accepted, is indexed by its position within the record. An existing index word is added to, or a new one is created. This often results in very large word indexes although this is not a disadvantage as modern software enables rapid retrieval from large bodies of text. Queries are built by selecting entries from the index which produce ‘hit lists’ of records found. Because free-text databases have traditionally not been good at handling numerical and structured data, hybrid systems have developed to maximise the functionality of free-text and relational structures.

Target-text databases: this type of data structure imposes query fields on to free text by tagging certain words and phrases. Extensive indexes are not produced as in free-text databases and data do not have to be rigorously structured as in relational databases. Instead, pieces of key information within free text are defined by tag names. The database software controls tag names and which tags can be searched on. The time and effort involved in tagging a document (using a markup language) is considerable although this can be semi-automated with editing software and the task is simplified if the text is structured within the document.

Text databases are important in many arts and humanities subjects and a lot of work has gone into the development and standardisation of text retrieval. SGML (Standard Generalised Markup Language) has emerged as the international standard for tagged-text databases Robinson 1994). SGML is also important for the electronic publication of documents as it can control layout and design. HMTL, the markup language for the World Wide Web, is a subset of SGML, and XML is an emerging new standard (see Crescioli et al. 2002 for an archaeological example). Here we can see the converging future between documents stored within text databases and the global access to hypertext documents offered by the Internet.

Despite its newness, XML has already had a significant impact on the working practices of a number of cultural heritage institutions, from the largest to the more modestly sized. The scalability and extensibility that are key features of XML fit in to the setup well at all shapes and sizes of organisations. As mentioned earlier in this report, a great many fields of research and business have their own custom-designed XML technologies, and an example is the Extensible Rights Management Language (XrML) which is used by content owners as a standard to make transactions regarding the use and provenance of cultural content..

In the cultural heritage sector, both SGML and XML have been used to great effect in archives and libraries, as well as in academic textual analysis – a discipline to which the concept and utility of markup have been familiar for decades. The Text Encoding Initiative (TEI) and the Encoded Archival Description (EAD) are two high-profile examples of organisations that have embraced Motto fulfil a number of functions, and to expand the range of the work they carry out. The uses of XML for digital preservation were explored in depth at an ERPANET workshop in Urbino, Italy in October 2002. The proceedings of this workshop are available online through the ERPANET website and provide a fascinating insight into the experiences of a number of European institutions.

5.1.1. Digital Asset Management Systems

The first DigiCULT Technology Watch (www.digicult.info) report gives a good introduction to Digital Asset Management Systems. Digital assets have the unique characteristic of being both product and asset. Some digital assets such as documents, images and Web pages are created and exist in digital form alone, while others like text, still images, video, and audio may be created through the digitisation of analogue material. Content has a value to institutions comparable to other assets such as facilities, products, and the less-measurable factor, expertise. Just as organisations have traditionally sought to maximise their use of financial, human, and natural resources, they will now aim to use digital assets to their full potential without having a negative effect on their 'value'. Value, of course, is not always obviously financial. Modern Digital Asset Management Systems (DAMS) provide mechanisms to manage digital resources. When associated or integrated with suitable policies, procedures, and licensing arrangements, DAMS provide a means of facilitating use of digital assets without impacting on the intrinsic value of the assets themselves. Digital Asset Management is the practice of using software applications and hardware (servers) to index, catalogue and store marketing tools and media assets in a centrally located, digital format. A DAM repository can be searched, shared, distributed and re-purposed to create a growing digital workflow environment, saving time and money, and achieving increased efficiency in communications.

DAM software is based around a central database or collection of linked databases, as is Customer Relationship Management software. Management and long-term preservation of the large volume of material likely to be held in a DAMS depend upon a storage management system capable of moving media entities between locations. An organisation's ability to harvest, reuse, and realise the value of its assets will only ever be as good as its mechanisms for storing and retrieving assets. DAMS can generally handle a number of different media types, such as creator-structured and XML documents, images, audio and virtual reality objects, and there are a number of systems which concentrate on media types for specific business areas. The broadcast media is a prominent example.

Matching a future DAM solution to existing database systems is of the utmost importance. It is also crucial to differentiate between the different types of database likely to be in use across an organisation. For example, if the majority of staff are currently working on standalone databases such as Microsoft Access, the information from these will have to be standardised and streamlined in order to facilitate distribution across a network. This is the only way that DAM software can fulfil its function efficiently. An upgrade to a more powerful client-server database will also be necessary if the assets are to be made available for public Web access. Object databases have been and are being developed which hold the assets themselves in a DBMS as user-defined objects. Storage of assets in a database will increase the granularity (or the level of detail at which the code or data is considered) of an organisation's assets to a significantly greater degree than the standard file approach. Storing assets in a file system with pointers to the location of the metadata can be more efficient, up to a point, especially when the content is linked with a freely-accessible Website. The decision on which strategy to adopt must depend on the size, formats and idiosyncrasies of the collections to be managed, as well as the organisation's ability to create and maintain an expensive single-purpose resource.

As is the case with many other technologies, the future of Digital Asset Management is heavily geared towards extensibility and XML compliance. New and developing products have the advantage of increasingly granular control, and facilitate the easy reuse of content. The primary advantage of XML-based systems is that they allow the separation of form and content, with the mark-up acting as a user-defined vessel for the content itself. Adobe's XMP (Extensible Metadata Platform) software allows metadata to be held in the header sections of individual files, permitting simultaneous transfer of content and custom-mapped metadata between discrete systems. In the manufacturer's words, XMP 'provides Adobe applications and workflow partners with a common XML framework that standardizes the creation, processing, and interchange of document metadata across publishing workflows.' The inclusion of metadata within the file structure allows the assets to retain their context even when accessed within a different application. XMP is based on the W3C's open standard for metadata, the Resource Description Framework (RDF), and is shared as an open-source license. Users and integrators are allowed free access to the source code via the Software Development Kit. Further developments in DAM are expected as a result of the new WebDAV standard, which allows collaborative authoring and editing of content via remote Web servers. The growth in object and object-relational databases will also have a marked effect on the DAM world, as database companies become more familiar with helping to define particular, sometimes company-specific and proprietary asset types.

5.1.2. Standards

The international and national information standards that museums and cultural heritage organizations require fall into four main groups:

- *Information system standards* define the functional components of the information system as a whole. For a museum, these might be the requirements for separate facilities in cataloging and collections management, membership, administration, finance, and publishing.

- *Data standards* define the structure, content, and values that collections information comprises. *Data structure* concerns what constitutes a record, such as the different fields used to record information and their relationships. *Data content* relates to the rules and conventions governing how data are entered into fields, including cataloging rules and syntax conventions. *Data value* has to do with the vocabulary used in the various fields and the specifications for individual character sets.
- *Procedural standards* define the scope of the documentation procedures needed to manage operations effectively. Examples include rules for logging on and off an electronic mail system, or policies governing an institution's acquisition or loan procdures.
- *Information interchange standards* define the technical framework for exchanging information, whether between systems in a single institution or among systems in multiple institutions. Examples of interchange formats include ISO 8879, Standard Generalized Markup Language (SGML); ISO 2709, originally developed to support the exchange of bibliographic information; and ISO 9735, Electronic Data Interchange For Administration, Commerce, and Transport (EDIFACT)--all developed by the International Organization for Standardization.

For a more extensive overview of standards in cultural heritage we refer to the EPOCH deliverable 4.2.1. (Report on Standards).

5.1.2.1. Museum standards

According to EMII standards "define procedures (a type of administrative metadata and units of information in a specific discipline in the cultural sector. The standards do not necessarily give guidance about its technical implementation. Procedural standards can be seen as defining 'good-practice' in a specific field. Descriptive standards focus on the definition of units of information. Sometimes procedural and descriptive standards are combined."

SPECTRUM – The UK Museum Documentation Standard, 2nd Edition[8]

Standard for the collections management documentation. Built around 20 procedures that commonly occur in museums. Supported by definitions of 'units of information' – the data needed to support the procedures.

CDWA – Categories for the Description of Works of Art[9]

The CDWA describe the content of art databases by articulating a conceptual framework for describing and accessing information about objects and images. They identify vocabulary resources and descriptive practices that will make information residing in diverse systems

[8] Published by mda (http://www.mda.org.uk)

[9] Published by the Getty Research Institute (http://www.getty.edu)

both more compatible and more accessible. They also provide a framework to which existing art information systems can be mapped and upon which new systems can be developed.

Dublin Core[10]

The Dublin Core is a simple metadata element set intended to facilitate discovery of electronic resources. Elements can be grouped into those having data on : Content – Coverage, Description, Type, Relation, Source, Subject, Title, Intellectual Property – Contributor, Creator, Publisher, Rights, Instantiation – Date, Format, Identifier, Language. Its use has been mandated by several governments in Europe (e.g. UK) and throughout the world (e.g. Australia). We include here Dublin Core though not directly related to museums for its wide diffusion in the heritage sector.

TEI – Guidelines for Electronic Text Encoding and Interchange[11]

Defines a set of tags (markers) for inserting into the electronic form of a document (a text) in order to identify the structure and other features of that document. The aim of these tags is to allow processing of the text by computer. They are aimed at use with texts in any 'natural language, of any data and of any genre. TEI is included here as the main text-encoding standard. However it is text-oriented and probably unsuitable for our goals.

5.1.2.2. Monument standards

MIDAS: A Manual and Data Standard for Monument Inventories[12]

MIDAS is an agreed statement of best practice for the compilation of inventories of monuments. It has been developed for all those who hold, or plan to develop, an inventory to record monuments, as heritage professionals, local authority managers, and so on.

5.1.2.3. Archaeological standards

Notwithstanding the claims by CIDOC-CRM supporters that there is an international standard, that is CIDOC-CRM, no such thing in fact exists. There exists, on the contrary, a plethora of national archaeological standards. Sometimes they are enforced by national regulations, as is the case in Italy or in the Netherlands. In several other cases there is no national provision of a 'standard' documentation system. Often regulations are described in the typical legal way, that is with reference/modifications to previous norms, and descriptions are difficult to find. In almost all cases, regulations are interpreted by individual archaeological researchers in the most diverse ways and lead to as many documentation standards as archaeologists, for instance by adding some field in forms, systematically not filling others, albeit prescribed, etc.

---

[10] Published by the Dublin Core Metadata Initiative (http://www.dublincore.org)

[11] Published by the Text Encoding Initiative Consortium (http://www.tei-c.org)

[12] Created and published by English Heritage (http://www.english-heritage.org.uk)

It must be noted that if one does not limit to Europe, there is an archaeological documentation system, which is expanding in the USA, called XSTAR[13]. IT has been developed at the Oriental Institute, University of Chicago. It seems that some kind of nation-wide endorsement or support is imminent; anyway, the system is now being used also outside the home institution.

5.1.3. Some database solutions for cultural heritage

IDEA - Integrated Database for Excavation Analysis[14] models the structure of archaeological data recording, rather than a specific data recording method, and then provides the user with a view of their data which corresponds with their method of data recording. IDEA is a relational DBMS written in Microsoft Access and provides a flexible, user-customisable framework for recording data from archaeological excavations. IDEA does all the hard work of establishing a 'clean' relational structure with around 80 tables, but it is in no way prescriptive about what is to be recorded or how it is to be recorded. IDEA accommodates archaeological concepts such as stratigraphic relationships, artefact 'lots', refitting sequences, reconstruction from fragments, hierarchical classifications, grouping of deposits into features, phasing etc., as well as many-to-many links such as those between artefacts, deposits, features and the drawings and photographs on which they appear. These are data structures which are often too complex to be incorporated directly into the structure of project-specific systems. IDEA handles all these structures without programming - relationships and classificatory schemes are built through menus and data entry forms.

Integrated Archaeological Database System[15], IADB's concept was formulated over five years ago. The project was initiated by Stephen Stead and has been developed by Michael Rains with archaeological support from Peter Clark and Richard Sermon. The IADB is built around the SUAT site recording system which itself has undergone development and refinement over time. The overall structure is essentially hierarchical and consists of five levels: Finds, Contexts, Sets, Groups, and Phases. The first two levels, Finds and Contexts, are formed by the Level II site records.

ArchaeoData[16] is an excavation database on Microsoft Access, developed since 1998 by ArcTron GmbH. This incorporates project management and finds processing, a scientifically defined finds book, and image processing, including the systematic recording of archaeological data. This database processes a wide ranging level of reports, as well as handling plans and profile sections. The description formulas are developed around various

[13] http://oi.chicago.edu/OI/PROJ/XSTAR/XSTAR.html

[14] http://acl.arts.usyd.edu.au/acl/products/software_utilities/idea/index.html

[15] http://www.suat.demon.co.uk/iadb/iadb.htm

[16] http://www.arctron.com/Software/ArchaeoDATA/index.php

archaeological criteria, such as pedology. The database can search and print out detailed archive material quickly via different adjustable data sourcing.

The Oxford ArchDigital[17] custom designs server- and client-based solutions that integrate databases, multimedia, and geographic information. Their solutions are built around a series of modules that function independently, or combined to form an extremely powerful Content Management System. Oxford ArchDigital builds for applications for all platforms (particularly Windows, Linux, and Unix) and wherever possible uses open source solutions including PHP, MySQL, and Apache.

Intrasis[18] (Intra-site Information System) is an archaeological information system for recording and managing field data. The system is based on GIS technology.Intrasis can be modified to suit many different kinds of excavations. There are basic functions built into the system to store and process data information. Several functions and extensions make Intrasis an powerful support for archaeologists.

M3 - minisis management for museums[19] is one of the worlds top cultural assets management applications. This tailored application for Museum Management created via the SMA (Standard MINISIS Application) enables members of the art, museum, archaeology and natural science communities to manage their mission critical tasks. At the same time, given the SMA toolkit, clients are permitted further customization to meet the changing demands of this ever-evolving industry.

ADLiB Museum[20] is an integrated Collections Management Application designed to support Spectrum procedures. The system is configurable to meet the requirements of individual Museums. ADLIB Museum includes integrated imaging and comprehensive terminology control facilities. Public access searching is available, with optional access through XML and ASP based web interface.

Cart@net[21], developed by Planetek Italia, is the solution for the management and consultation of large raster datasets, ideal to distribute on-line catalogs of cartographic data. Cart@net allows to visualize and interrogate a cartographic vector database by Internet, effecting researches based on graphic and alphanumeric standards, and to visualize the raster cartography related to the area of interest. In fact the whole raster database is accessible, without solution of continuity, together with the window of the vector data. The system, developed to integrate with vector servers, results perfectly integrated with both Autodesk

[17] http://www.oxarchdigital.com/index.php?page=5&expand=7

[18] http://www.raa.se/uv/intrasis/eng/intr_about01.htm

[19] http://www.minisisinc.com/index.php?page=m3

[20] http://www.adlibsoft.com/(ucnqgm55lrot1055lzyrr455)/default.aspx?language=en&office=UK

[21] http://www.planetek.it/eng/cartanet.asp

Map Guides, server for the distribution of vector cartography, that allows to create, publish and distribute extremely detailed vector maps on Intranet and Internet, and Image Web Server, for the distribution of raster data.

5.2. Future directions and EPOCH's contribution

Although storage and retrieval of multimedia information in and from databases are the focus of other 6FP initiatives, it is nevertheless useful to mention an aspect that has not been considered intensively in the context of Cultural Heritage, namely the retrieval of images not based on textual annotations (for this see e.g. the elegant framework 6FP Integrated Project BRICKS is developing), but on the direct analysis of their content. We believe that such technologies will become important instruments for scholars and the public alike, when the technologies come at age. One can easily imagine a host of applications, like detecting stolen pieces of art that are offered for sale, finding similar types of artefacts in musea or found at excavations, mobile museum guides that automatically download information on any exhibit that a camera or mobile phone is pointed at automatic detection and monitoring of archaeological sites from aerial images, etc.

Although content-based image retrieval is not a focus of EPOCH for the moment, the network feels it has expertise among its partners that ought to be harnessed also in the realm of Cultural Heritage in the longer run.

5.2.1. Progress and projects

Current approaches to content based image retrieval are still quite primitive. There are currently two main approaches to producing indexes for video search:

- Manual annotation still predominates for content description in TV archives, stock footage agencies, and film libraries. Except for the International Press and Television Archive (IPTA) guidelines for news data, there is little standardisation between sites. Most thesauri are kept in proprietary formats that have been evolved in an ad hoc manner over a number of years. Manual annotation is meticulous, costly, entrenched within organisations, and has not been considerate of automation.

- For certain genres like TV news where the spoken word bears much of the content, it is reasonably effective to index using text generated from the speech (e.g.\ through automated speech recognition). However this does not extend to genres where speech is absent or less directly informative, such as documentary videos or collections of stills. Text-based indexing is also of little avail when the search is for incidental data that were not the focus when the footage was originally taken, e.g.\ a building visible in the background that was recently destroyed by a bomb blast. In all of these application domains, it is essential to index on visual information directly.

Recent announcements by Google Inc. and by Yahoo~! have pointed to those companies launching so-called image and video search engines. These are, in fact, no more than text searching through closed captions with pictures or still images (keyframes, taken from the video) added for presentation. Hence, this is very different to using the actual image content.

The use of actual image content has been quite restricted up to now, and efforts have focused on stills rather than video, often focusing on keyframes if video must be analyzed. In the area of image based retrieval, several systems are on the market. One of the most widely used is IBM's QBIC \cite{IBM}, which features in products such as IBM's DB2 Digital Library, as part of the DB2 Image Extenders (components of IBM's scalable, multimedia, Web-enabled DB2 Universal Database). It offers retrieval by the combination of colour, texture, and shape. But shape is related to clean object outlines, which only works well with uncluttered backgrounds, and IBM themselves put much more emphasis on retrieval based on color and texture. Another system is Virage , e.g. used in AltaVista's AV Photo Finder. For the analysis of video data, the Videologger product has been released. Again, cues for retrieval are kept quite basic, such as color and texture, as well as motion. Convera (VisualRetrievalWare) is a similar product, with the Yahoo! Image Surfer as its best-known application.

Besides these commercial systems, several experimental research systems have been developed. Photobook from MIT was an early example, that supported retrieval from textures, shapes, and human faces. Chabot, later renamed `Cypress', provided a combination of text-based and colour-based queries and was incorporated into the Berkeley Digital Library project. Cypress' CBIR capacity seems to have been taken out more recently, but Berkeley has continued such research with their Blobworld software. It tries to segment objects out of the scenes, which is very hard. Recent work at Berkeley detects specific object classes like horses and naked people. But the success rates of this segmentation based object class detection are still too low for practical purposes. VisualSEEK was the first of a series of systems developed at Columbia University. A strong point was that queries were quite flexible in the spatial layout specifications that they could accept. WebSEEK was a further development, aimed at easing queries over the Web. Emphasis is on colour, but in another prototype called VideoQ motion was added. Relative layouts of regions with specific colors or textures can now also be specified. However, most object classes cannot be detected reliably using such simple layouts. The MARS project at the university of Illinois put emphasis on control by the user with its relevance feedback mechanisms, as did NEC's PicHunter. An example of European CBIR technology is the Surfimage system from INRIA. It has a similar philosophy to MARS, using multiple types of image features that can be combined in different ways, and offering sophisticated relevance feedback. The RSIA system developed by ETH Zurich and DLR is dedicated to search in large satellite image databases. It supports remote access through the Web and is based on texture features. The feature set does not include the more geometrical kinds of structures that would be needed for object recognition (e.g.\ finding archaeological sites from such imagery). The Netra system uses similar cues to many of the above systems --- colour, texture, shape and spatial location information --- but as in BlobWorld, a more sophisticated segmentation function is included. Synapse is a rather different kind of system, in that it bases its similarity judgements on whole image matching. This list of non-commercial image retrieval systems or projects is far from complete, of course. Other examples are the Picasso, IQuest, Compass, SIMPLIcity, ImageRover, Imedia, Viper, ImageScape, ImageMiner (formerly IRIS), and PicSOM systems. They tend to be more focused on particular subproblems.

FP5 projects related to content-based image retrieval included the query-oriented AMICITIA and POP-EYE, the ontology-oriented MUMIS, the content negotiation agent oriented DICEMAN, and several image processing oriented projects including ADVISOR (refined shot deteection and grouping tools) an VICAR (combining low-level features for appearance-based video search). VICAR's Video Explorer was a step beyond the traditional histogram style approaches, but it was limited to objects that always occur with the same layout of basic features and did not attempt semantic-level search.

In FP6 there are a number of funded projects relating to video content, but most of these, including HUMAINE, REVEAL THIS, Polymnia and others, concentrate on using low-level video features for content description and matching rather than targeting high level semantic information (although an onset can be seen in Reveal-This). FP6 IP aceMedia targets analysis, indexing and content manipulation of video content. Yet, further efforts are needed towards fully automatic content analysis to support efficient indexing on fine-grained semantic-level video entities (aceMedia is oriented more towards ontology construction and using existing annotations and low-level features for clip-level retrieval).

A good overview with links to related projects can be found at http://www.cdvp.dcu.ie/links.html.

In summary, colour and texture have dominated image content search. In the rare cases where objects are retrieved, the methods assume that their outlines or parts can be found through segmentation procedures, which are notoriously fragile. Also, in almost all cases (except for some work on faces and the work at Berkeley), specific objects can be identified, but not object classes that show substantial within-class variability. It is fair to say that the state of the art in visual content indexing and navigation is well behind the current sophistication of content operations on text, where one can search, link, cluster, summarise, extract facts, answer questions, and perform cross-lingual operations.

5.2.2. EPOCH

Some EPOCH partners contributed directly to a number FP5 projects that were among the first to change this state of affairs. Foundations for object detection in cluttered scenes have been laid in VIBES (Video Browsing, Editing and Structuring) and CIMWOS (Combined IMage and Word Spotting). Initial work on action recognition (single persons) was also carried out in VIBES. A sound substrate on which to build large-scale object categorisation was generated in CogViSys (Cognitive Vision Systems), and in LAVA (Learning for Adaptable Visual Assistants), two FP5 Cognitive Vision projects not focusing on image retrieval, but nevertheless looking into the relevant issues of feature selection and model building for object classes. CIMWOS was among the first projects to take a multi-modal point of view, combining text, speech, and image content, depending on their combined availability.

Further information on content-based image retrieval can be found in (Rui 1999, Goodrum 2000, Veltkamp 2002 and Venters 2001.

5.3. References

Convera: http://www.convera.com/

Crescioli, M., D'Andrea, A. & Niccolucci, F. 2002. XML encoding of archaeological unstructured data. *Proceedings of CAA 2001*: 267-275.

DigiCULT TechWatch Report 1. February 2003 http://www.digicult.info..

Goodrum, A., Image Information Retrieval: an overview of current research, *Information Science*, Vol.3, No.2, pp.~63-66, febr. 2000

IBM: http://www.qbic.almaden.ibm.com

Lock, G. 2003. *Using computers in archaeology. Towards virtual pasts*.

Robinson, P. 1994. *The transcription of Primary Textual Sources using SGML*. Oxford.

Rui, Y., T. Huang, and S.-F. Chang, Image Retrieval: current techniques, promising directions, and open issues, *Int. J. of Visual Communication and Image Representation*, Vol.~10, pp.~1-23, 1999

Smith, D. 1991. Database fundamentals for archaeologists. In S. Ross et al. (eds): *Computing for Archaeologists*: 111-125.

Veltcamp, R. and M.Tanase, *Content-based image retrieval systems: a survey*, Technical Report Un. of Utrecht, oct. 2002 (http://www.aa-lab.cs.uu.nl/cbirsurvey/)

Venters, C., R. Hartley, R. Cooper, and M. Hewitt, Query by visual example: assessing the usability of content-based image retrieval system user interfaces, *Proc. 2nd Int. Symp. on Multimedia Information Processing*, 24-26 oct. 2001, Beijing

Virage: http://meru.cecs.missouri.edu/mm_seminar/virage.html

6. Mobile wearable and ambient systems

6.1. Systems, tools and technologies

6.1.1. Mobile

Here, we are concerned with technologies that support, monitor, or adapt to, the mobility of people, artefacts, devices and software at all stages of the pipeline from data acquisition to public presentation.

Small, portable devices are increasingly used in field data collection as well as education and training in field methods and skills. The use of such devices can be increasingly 'connected' as it becomes relatively easy to deploy networked and distributed data collection and management systems during field campaigns. Within museums and other cultural institutions, artefact tracking might be employed as an element of collections management. For example, using RFID or optical tags to aid the tracking of objects between storage and display, or through conservation processes. Additionally, sensor devices might be used to monitor the environmental conditions under which an object is stored or transported. In the public presentation of museums and sites we can envisage a steady development beyond existing audio guides using small portable multimedia devices that range from personal audio and image players, through mobile smart phones, PDAs and other, more specialised, but minimally intrusive devices.

6.1.2. Wearable

Wearable computers are typically somewhat larger and heavier, and are usually carried on a belt or in a backpack, or suspended from neck or shoulder straps. In some cases, they may be divided into a number of communicating components and distributed over the body to improve weight distribution or heat dissipation. The aim of such systems is to provide a desktop or laptop level of computing resource without significantly impeding the user's movement, or ability to perform tasks. Conventional input devices are typically replaced by wearable equivalents ranging from single handed chording keyboards and 3D mice to keyboards woven into smart fabrics. Output frequently involves head mounted displays, ranging from lightweight overlay systems worn like spectacles to fully-immersive VR headsets.

6.1.3. Ambient

The research area variously known as Ubiquitous Computing, Pervasive Computing or Ambient Intelligence derives much of its inspiration from Weiser's vision of a third age of computing beyond the current 'Personal Computing' paradigm and the growth of the Internet. The vision is characterised by the ubiquitous presence of networked computing devices, on the person, in vehicles, in the fabric of buildings, in consumer products, etc. We are already some way towards this with embedded processors and mobile phones greatly outnumbering conventional computing devices but, in our present environment, only a minority of these devices are networked and even fewer are more than minimally interoperable. In this imagined future, we will interact directly with only a small proportion of the devices around us. Unlike today, where the computer is often the centre of attraction, many of these devices will be peripheral and will disappear into the environment where they will provide information, services and control functions as and when they are needed. In this paradigm, a personal computing environment should be truly personal in that it accompanies the individual wherever they go and whatever they are doing. It should not, however, be limited to the capabilities of our conventional desktop or laptop machines. Instead, it should also be able to adapt to its immediate environment and to make use of location specific services. Equally, the environment should be able to adapt to its occupants by offering services that most closely correlate with their needs.

6.2. Future directions and EPOCH's contribution

For many reasons, applications in the domains of cultural and natural heritage have proved to be especially popular amongst ubicomp researchers as a way of demonstrating innovative ideas. Amongst these reasons are the inherent mobility of potential users, a wide diversity of attractive materials for presentation, and the potentials of tourism and associated markets. Nevertheless, it must be emphasised that most of what has been achieved so far has resulted in one-off demonstrators, and we are still looking forward to a period when deployment of interoperable systems becomes a reality.

To date, most demonstrator applications in the cultural and natural heritage areas have concentrated on data collection systems, including their use in education, and various forms of visitor guides. These have been seen as obvious applications of context-aware and, more specifically, location aware technologies.
Location-awareness has been an especially popular subarea of context-awareness. Different location technologies are appropriate to indoor and outdoor location. Indoors, these are mainly short-range infrared, ultrasonic or radio signals for indoor applications, such as museum guides. Outdoors, Global Positioning System (GPS) receivers have been widely used

for data collection and urban visitor guides. Other, typically lower precision, approaches include those exploiting the infrastructure of mobile phone networks and the increasing ubiquity of WiFi base stations. Many examples using location as a key aspect of context illustrate the idea of smart environments in which information and services applicable to immediate local needs are emphasised.

In most such applications, the locations of interest are typically those of the mobile user and of static exhibits and structures. The location of portable artifacts has received less attention, but there is clearly scope for applications that support the specialist analysis of collected data, and the management of museum collections.

A key ingredient in building context and location-aware services for smart environments is an infrastructure for managing and disseminating contextual information. It is now possible to envisage the deployment of such infrastructure support for the full spectrum of activities from initial data collection, through the essential analytical, archival and curatorial stages, to the end products of public and scholarly presentation, education and edutainment.

Applications need not be restricted to an individual museum or visitor site. The concept of 'cultural routes' (recently introduced by Neil Silberman and Daniel Pletinckx of the Ename Center for Public Archaeology and Heritage Presentation, Ename, Belgium) seeks to bring together related attractions within a region and to offer thematic routes that would be attractive to individuals and family groups pursuing self-organised vacations. This, in turn, adds complexity to location models, and suggests a link with more conventional navigational tools.

This wider range of scales brings with it a need for careful consideration of established ubicomp concerns, including

- networking and communications
- the suitability of different devices for use in a range of situations
- interoperability of devices and infrastructures
- support for tool building
- support for content authoring
- and, of course, personal privacy.

6.2.1. Networking and Communications

Addressing communications over the range of scales envisaged in the previous section will require different scales of communication. This suggests that devices should be able to adapt transparently to the available networks, taking account of Quality of Service offered whilst minimizing costs to the user.

Reliability and availability of connection is also a well-known issue. When data collectors work in remote locations, mobile phone networks are invariably the only available form, yet even these have intermittent coverage of many mountainous areas and locations far from population centres. Similar problems may arise, albeit less frequently, for those following cultural routes which pass through remote rural areas.

Even indoors, where WiFi or Bluetooth connectivity is available, coverage may be partial or bandwidth may be reduced as the number of participants increases. The problem may be particularly acute in modern buildings containing large amounts of structural steel.
In summary, heritage applications will need to address the full range of connectivity and communications issues that have been recognized by the ubicomp community.

6.2.2. Devices

Whether a general-purpose "life navigator" accompanying all of our mobile experiences may ever really exist is still an open question. Certainly current standard devices are often inadequate. For example, in the cultural heritage domain, it is doubtful that standard PDAs can be successfully used as museum navigators, for the following main reasons:

- Their lack of embedded context-management support
- Their inability to be accessed hands-free
- Their internetworking bandwidth
- Their screen size and brightness

Better solutions may be provided by application-specific navigators with ergonomics optimised to meet the user requirements, and with functional specifications inspired by research in the following areas:

- Context management and activity recognition
- Context-based usability policies
- Context-based resource and power management
- Space modeling
- Technology convergence

6.2.3. Interoperability

So far, most heritage-oriented systems have typically been based around a single class of user terminal device, such as a production PDA. In a few cases, specialized devices have been developed to address limitations of commercial products. However, looking to the future, we must be prepared to deliver services to a variety of different devices, including special purpose terminals that are beginning to replace an earlier generation of audio guides, but also to PDA-like machines and, especially, smart phones owned by the visitors.
In the cultural route model, there is also a requirement for interoperability with vehicle and personal navigation systems, as well as across and between support infrastructures. These infrastructures may be commissioned and operated by different commercial or governmental institutions, yet visitors will need to move smoothly between their coverages. There are many open issues of importance to the ubicomp community concerning how this might be achieved with a minimal impact on the developers, implementers and managers of these separate systems. Ideally we should be able to build once, yet interoperate with many.

6.2.4. Tool Building

The development of applications based on smart environments requires a considerable integration effort, and suitable tools are required to support both application set up and algorithms verification.

For example, in the cultural and natural heritage domain we need to devise new sensory solutions in order to recognize the users activity and their focus of attention, with the goal of anticipating their intentions; research on sensor systems for wearable devices should be carried out, the main goal being to close the gaps — that is cancelling all discontinuities — between the visitor, the environment and the platform. Similar challenges may have to be considered in other domains.

6.2.5. Content Authoring

Usually, when we access a medium, two actors are involved: us and the medium itself (e.g. the computer, the stage, the television screen, the newspaper). In contrast, at a cultural or natural heritage site, and more generally in most mobile context-related applications, there are three actors: ourselves, the mobile device and the "target" of our attention, i.e. an application-specific real object (for example the exhibit in museum applications). Multimedia content, therefore, should not surrogate the target real object, as normally happens in a DVD or on the Internet, but should be the catalyst of the "resonance" between the users and their environment. So, we need to find new cognitive models, new ways of mixing audio, video and text, new methods to handle interactivity and to create "contextual multimedia" for on-site access.

This issue is already considered by many cultural institutions. The contents currently under development within the MUSE project, for example, are based on research carried out at two Italian museums, the Certosa e Museo di San Martino in Napoli, and il Museo di Storia della Scienza in Florence. But, a similar approach could be applied to other applications where location-based services are involved. The time has come to investigate content authoring methodologies for location-based services and the experience gained in cultural and natural heritage applications may have the potential to be used in more general frameworks.

6.2.6. Personal Privacy

Privacy is a recognized issue in ubiquitous computing, albeit one that has received relative little attention. In the kinds of systems envisaged here, it will be an important consideration where many users will be expected to visit multiple sites/museums and where there may be benefits to the user experience to be gained by sharing information about the user's activities and interests between sites. Potentially there may be conflicts between personal privacy and commercial interests as different sites may be operated by separate companies who may view visitor information as commercially confidential.

6.2.7. Available experience within EPOCH

Several EPOCH partners have valuable experience in one or more of these technologies and their applications resulting from past and ongoing projects. In the Data Capture area, mobile data acquisition tools are represented by the RAMSES project (partner 31 DISI) which focused on excavation recording, and the ongoing FieldMap project (partner 52 UNIKENT, field evaluation in collaboration with partner 84 RUG) focussing on field survey. Spanning the areas of Data Capture and Interpretation and Analysis is *jnet* (partner 52 UNIKENT), a graph visualisation and manipulation tool for stratigraphic data from excavations. A mobile version of *jnet* has been developed from an earlier prototype] as part of EPOCH showcase *Tools for Stratigraphic Data Recording*.

The other main area in which EPOCH partners have relevant experience is that of Public Presentation. Here, again, partner 31 (DISI) played a prominent role in the PAST project which sought to aid public understanding of excavated sites. Another Italian group (partners 30, ARCES; 32, DS; 29, POLIMI) were responsible for the MUSE project and the development of a wearable terminal device WHYRE. Elsewhere, other partners (24 INTRACOM, 20 FHG-IGD) have been major contributors to the ARCHEOGUIDE project and subsequently have demonstrated a lighter weight augmented reality system on handheld devices.

6.2.8. EPOCH NEWTON project CIMAD

Several of the key issues covered in the previous section will be addressed in the EPOCH NEWTON project CIMAD ({Common Infrastructure, Context Influenced} Mobile Acquisition and Delivery of CH data).

The main goal of CIMAD is the exploratory implementation of a framework for smart CH environments supporting distributed and mobile on-site applications, from data capture to public dissemination. The framework is intended to become a component of the EPOCH Common Infrastructure. Within CIMAD a Smart Environment is any defined area of cultural interest (for example a Museum, a Temple, an Archaeological Site, an historical City Center, an industrial heritage site, an historical cemetery...) where there are means of enabling devices to detect the context of human participants, so that contextual information can be used to support and enhance their abilities in executing application specific actions. These means may be sensor systems embedded in the environment itself, sensors integrated in the platform or both. The smart environments generated with the proposed framework will include networked context-aware mobile devices, repositories, context-servers and stationary clients.

The proposed framework will include three modules:

- an Application Builder,
- a Context Management Infrastructure,
- Authoring procedures for context-aware mobile applications.

CIMAD is a collaborative project involving the following EPOCH partners:

- 30 ARCES: Advanced Research Center on Electronic Systems for Information and Communication Technologies "Ercole De Castro", Università di Bologna, Italy.
- 52 UNIKENT: Computing Laboratory, University of Kent, Canterbury, UK.
- 29 POLIMI: HOC (Hypermedia Open Center), Department of Electronics and Information, Politecnico di Milano, Italy.
- 32 DS: Ducati Sistemi S.p.A., Bologna, Italy.
- 68 IBC: Istituto per i Beni Artistica, Culturali e Naturali della Regione Emilia Romagna, Bologna, Italy.

6.3. References

Abowd G. D., Atkeson C. G., Hong J., Long S., Kooper R. and Pinkerton M., Cyberguide: A Mobile Context-Aware Tour Guide, *Wireless Networks*, 3(5), 421-433, 1997.

Ancona, M., G.Dodero, V.Gianuzzi, O.Bocchini, A.Vezzoso, A.Traverso, and E.Antonacci, Exploiting wireless networks for virtual archaeology: the PAST project, in F. Niccolucci (ed.), *Virtual Archaeology, Proc. VAST EuroConference,* Arezzo, Italy, November 24–25, 2000. BAR International Series 1075, pp. 195-199.

Ancona, M., G. Dodero, M. Mongiardino, A.Traverso, Taking digital Notes in the Field: the Archeo tool-set, in J.A. Barcelo, I. Briz and A. Vila (eds), *New Techniques for Old Times: Computer Applications in Archaeology, 1998*, (BAR S755), Archaeopress, Oxford 1999.

Davies, N., Mitchell, K., Cheverst, K. and Friaday, A., 1999, Caches in the Air: Disseminating Tourist Information in the Guide System, *2nd IEEE Workshop on Mobile Computing Systems and Applications*, New Orleans, USA, 11-19.

Galani, A., and Chalmers, M., 'Can you see me? Exploring co-visting between physical and virtual visitors', *Museums and the Web 2002*, Boston, US, Archives and Museum Informatics.

Kostov V., Ozawa J., Matsuura S., Analysis of Wearable Interface Factors for Appropriate Information Notification, *8th IEEE International Symposium on Wearable Computers (ISWC2004,* Arlinton, Virginia), Oct 31 – Nov 3, 2004, Proceedings, pp. 102 - 109

LaMarca, A. *et al.* Place Lab: Device Positioning Using Radio Beacons in the Wild, Intel Research, IRS-TR-04-016, October 2004. http://placelab.org/publications/pubs/IRS-TR-04-016.pdf

Ma L., Smith D.J., Ryan N.S., Acoustic Environment as an Indicator of Social and Physical Context, *to be published in Personal and Ubiquitous Computing, 2005.*

Pascoe, J., Morse, DR., and Ryan, NS., Developing personal technology for the field, *Personal Technologies*, 2:28-36, August 1998.

RAMSES project: http://www.disi.unige.it/person/DoderoG/ramses/main.htm

Rogers, Y., Price, S., Randell, C., Stanton Fraser, D., Weal, M. and Fitzpatrick, G., 'Ubi-learning Integrates Indoor and Outdoor Experiences', *CACM*, 48, 1, January 2005.

Ryan, N. 2001. Jnet: a successor to gnet. In W. Borner, editor, Archaologie und Computer, Workshop 6. Forschungsgesellschaft Wiener Stadtarchaologie, November 2001

Ryan, N. S., (2002), ‘Back to Reality: Augmented Reality from Field Survey to Tourist Guide’, in F. Niccolucci (ed.), *Virtual Archaeology, Proc VAST EuroConference*, Arezzo, Italy, November 24–25, 2000, 45-52, Archaeopress, Oxford, 2002

Ryan, N.S., Smart Environments for Cultural Heritage, *24th International Research Symposium, International Research Center for Japanese Studies*, Kyoto, February 7-11, 2005. http://www.cs.kent.ac.uk/~nsr/papers/kyoto2005/SmartEnvCH.pdf

Ryan, N.S., Pascoe, J. and Morse, D.R., Fieldnote: a handheld information system for the field, in R.Laurini, (ed), *Proc. TeleGeo'99, 1st International Workshop on TeleGeoProcessing*, 156-163. Claude Bernard University, Lyon, 1999.

Salmon Cinotti T., Nagaraj R., Mincolelli G., Sforza F., Raffa G., Roffia L., WHYRE: a Context-aware Wearable Computer for Museums and Archaeological Sites, *8th IEEE International Symposium on Wearable Computers (ISWC2004,* Arlinton, Virginia), Oct 31 – Nov 3, 2004, Proceedings, pp. 174 - 175

Salmon Cinotti, T., Raffa, G., Roffia, L., Garzotto, F., Muzii, R., Varlese, V., Malavasi, M. and Galasso, S., Evaluating Context-aware Mobile Applications in Museums: Experiences from the MUSE project, *Museums and the Web, 2004*, http://www.archimuse.com/mw2004/papers/salmon/salmon.html

Spasojevic, M. and Kindberg, T., A study of an Augmented Museum Experience, HP Laboratories, Palo Alto, July 2001, http://www.hpl.hp.com/techreports/2001/HPL-2001-178.pdf

van Leusen, M. and Ryan, N.S., Educating the digital fieldwork assistant. In G. Burenhult, G. (ed), *CAA 2001: Proceedings of Computer Applications and Quantitive Methods in Archaeology Conference*. Gotland University, Visby, Gotland - Sweden, Archaeopress, Oxford, April 2001

Vlahakis, V., Demiris, A., Bounos, E. and Ioannidis, N., A novel Approach to Context-Sensitive Guided e-Tours in Cultural Sites: “Light” Augmented Reality on PDAs, in Chrysanthou, Y., Cain, K., Silberman, N. and Niccolucci, F. (eds), *VAST 2004, 5th International Symposium on Virtual Reality, Archaeology and Intelligent Cultural Heritage*, December 2004

Vlahakis, V., J. Karigiannis, M. Tsotros, M. Gounaris, L. Almeida, D. Stricker, T. Gleue, I. Christou, R. Carlucci, N. Ioannidis. "ARCHEOGUIDE: First results of an Augmented Reality, Mobile Computing System in Cultural Heritage Sites", *Virtual Reality, Archaeology, and Cultural Heritage International Symposium (VAST01)*, Glyfada, Nr Athens, Greece, 28-30 November 2001.

Weiser, M. and Seely Brown, J., The Coming Age of Calm Technology. In Denning & Metcalfe (Eds.). *Beyond Calculation: The Next 50 Years of Computing*. Springer-Verlag, 1997.

7. Recording and data presentation

7.1. Systems, tools and technologies

7.1.1. Overview of 3D acquisition techniques

There is a wealth of 3D acquisition techniques. Yet, not all are equally relevant for cultural heritage, where the need for high precision is typically quite outspoken. This renders many of the so-called `shape-from-X' techniques less appropriate, for instance. Shape-from-silhouettes is one such technique. An object is put on a turntable and a camera takes images at equal, angular intervals. The object is modelled as the intersection of the cones that are formed by considering all projection rays though the object's silhouettes in each view. Such approach is very simple and elegant, but cannot capture cavities. Several other 3D acquisition techniques have similar limitations or are simply insufficiently robust (shape-from-shading, shape-from-texture, shape-from-symmetry, photometric stereo, etc.). The remaining, dominant technologies can be subdivided into two categories: *time-of-flight* scanners and *triangulation based* systems.

As to the former, the basic principle is to send out a wave or pulse, and to measure the phase shift or time delay resp., of the re-entering signal, reflected by the point in the scene the signal was directed to. By choosing a sequence of such directions, the surrounding scene is scanned.

Triangulation-based systems locate points in space as the intersection of different rays. These can be the two projection rays of a point as captured with two cameras, but also the intersection of a laser ray and a camera projection ray. Indeed, finding the projections of the same point in two images may seem trivial to a human observer, but is actually hard to do automatically. This is all the more the case if the surface of an object is homogeneous (i.e. `untextured'). Therefore, the second camera is often replaced by a laser, which projects a point onto the surface. It is then relatively easy to find this point in the image of the remaining camera. When only cameras and no special illumination is used, the method is referred to as `*passive*'. In case special illumination like a laser is used, the method is called `*active*'. Next, recent developments in both types of systems are discussed.

Next, we discuss some variations and new developments in passive, triangulation based systems.

Rather than trying to determine 3D shapes from a pair of images (`stereo'), one can use several. The difference doesn't only lie in the number of images and the better quality and completeness of the resulting models, however. When using stereo, the relative positions of the cameras as well as their parameter settings (e.g. focal length) have to be determined in order to calculate the equations of the projection rays that are to be intersected. Such parameter determination is called *calibration*. Calibration tends to be a rather tedious procedure, with special calibration patterns that needed to be positioned at specific places with respect to the stereo system. The more flexible the stereo setup is, the more often calibration typically is called for. If more than two images are available (i.e. if more than two camera positions are combined), then such procedures may be unnecessary. Indeed, working with multiple images offers the possibility to let the computer figure out what the parameters of the cameras are, directly from the images taken to model an object or a scene. Such procedure is called `self-calibration'. Recent systems therefore allow the user to move around with a hand-held camera, take images, and automatically get 3D data out if the images cover the surface sufficiently densely and the camera has been moving along an appropriate path (no pure translation or rotation). This turns consumer hardware like a handheld video camera into a 3D acquisition device. The advantages are obvious: low cost and large flexibility. The self-calibration techniques have been largely developed by European research teams, in the scope of IST 3FP and 4FP projects like VIVA, CUMULI, and VANGUARD, and have lead to several spin-off companies (2d3, Eyetronics, Realviz) offering such solutions. Such `Shape-from-video' techniques have been used in the context of cultural heritage by the 5FP 3D-MURALE project.

Basically, shape-from-video is a software solution based on a single camera that is moved around, whereas stereo systems tend to come as a dedicated hardware setup. Usually they include a rig with two mounted, fixed cameras or two cameras within the same housing. An advantage of a stereo setup over shape-from-video is that it can better handle dynamic scenes. Indeed, in a stereo setup the two images are taken simultaneously and hence it is as if the scene would have been static. But when the images are taken subsequently, with a moving camera, motions in the scene may cause problems (trees moving in the wind, people walking through the scene, etc.). Another advantage is that stereo images can be taken with still cameras, which tend to have a higher resolution than video cameras. Hence, recently, shape-from-video methods have been applied to multiple, still images as well. The main difficulty there is that point correspondences have to be determined between images that are more different than subsequent video frames. This problem, which such methods share with wide baseline stereo systems, has been solved to a substantial extent in recent developments.

Passive, triangulation based systems basically measure the difference in parallax, i.e. shifts of a point's projection in different images, for different distances. Such parallax can only be measured at the precision of a pixel or maybe 1/10 of a pixel at best. Hence, it is important that the parallax effects extend over several pixels, if the measurements are to be made at a sufficient precision. As parallax quickly goes down with distance points that are further away from the cameras will be measured will less precision. Precision can be increased by increasing the parallax, which can often be achieved by taking images from positions that are farther apart. Of course, the point should remain visible in multiple images. If a system is to be flexible in the sizes of objects it can deal with, then one should be able to adapt the camera distances, i.e. the *baseline*. This is mainly important for stereo systems, where only two camera positions are considered. Therefore, stereo cameras within a single housing will never do as a 3D capturing system at larger distances. On the other hand, when 3D reconstructions are produced on the basis of a moving camera, the overall distance that the camera has travelled will determine how precise measurements at larger distances are. The right half of figure 5 shows uncertainties that exist on the locations of points. As can be seen, this uncertainty goes up quadratically with distance from the position of the stereo system.

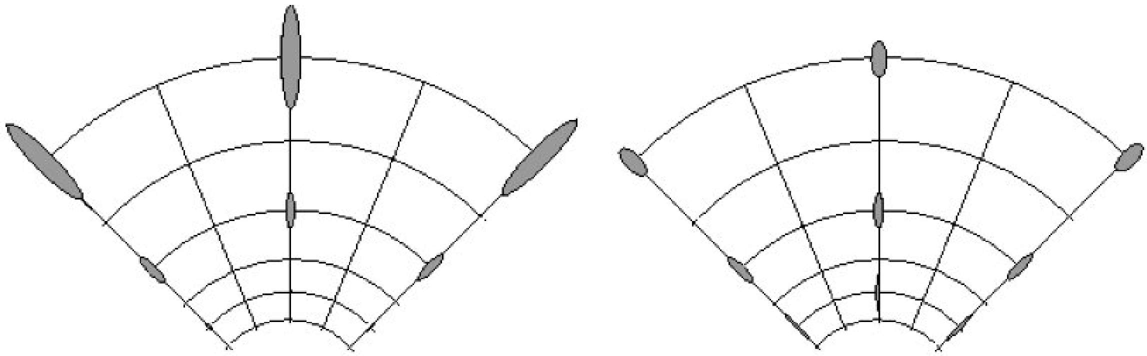

Figure 5: Precision of measurements with a stereo system

Time-of-flight systems show a very different pattern of errors. Here the precision depends on the precision with which time differences between emitted and reflected signals can be measured. As long as the intensity of the returned signal is high enough to be detected reliably, the resulting error is basically constant. This is shown in the left sketch of Fig. 5. Errors typically are a couple of centimetres. Comparing these error profiles, it can be seen that there is a cross-over distance, beyond which time-of-flight errors get smaller than those of a stereo systems with a fixed baseline. As this baseline would be increased, this distance would become larger. Increasing the baseline of a stereo system will eliminate ever larger parts of the scene close to the system as points there will not be visible to both cameras. Shape-from-video has the advantage that many images can be taken and points can be visible in multiple images where there always are a couple taken from vantage points that are sufficiently far apart.

Self-calibrating, passive systems may seem the holy grail of 3D acquisition, but will not always be the best option. The problem with untextured surfaces remains. Hence, there is a clear need for active systems like time-of-flight or active triangulation systems as well. As the laser needs to be directed towards all points to be located in 3D, this may become a slow and expensive solution. Rather slow capture is and remains a limitation of time-of-flight systems. In the case of active, triangulation based systems, acquisition speeds have increased substantially over the years. For instance, the laser point can be spread out into a line, using a cylindrical lens. This reduces the necessary 2D scan motion to a much simpler 1D sweep. More recently, scanning motions have been eliminated altogether. Rather, series of 2D patterns are projected, one after the other, and each time an image is taken. Together, the observed patterns identify the projector ray by which the different points are hit. The patterns encode this ray. There is a trade-off between the number of projected patterns and the complexity of the decoding. One-shot systems represent a limit case, where only a single pattern is projected, and the decoding takes quite a bit of computations. Yet, given the ever increasing speed of computers, one-shot systems may still be equally fast or even faster. And they acquire all the necessary data from a single image, which makes it possible to take such a system in the hand while acquiring the data. Such portable solutions add flexibility to the acquisition process and reduce the overall times needed to build complete object models.

In summary, recent years have witnessed enormous progress in the flexibility and reduction of cost of 3D acquisition systems. Yet, all solutions still have important restrictions. Time-of-flight systems are quite slow, don't yield a surface texture, show rather large relative errors over short distances, and work from a point, which renders them much more apt for the capturing of a large environment than of smaller objects. Passive, triangulation based systems have evolved into very flexible solutions, that can work with nothing but a handheld camera. Textures come naturally from the same images that are used for the 3D acquisition. On the other hand, there has to be sufficient texture on the surfaces for them to be captured successfully. Active, triangulation based systems are less flexibly, in that a projector has to fixed to the camera. Although small and lightweight solutions have been proposed, these then capture rather small parts of the surface at the time, typically without the texture. Also, the working volume / distance of such systems tends to be specific and limited.

7.1.1.2. System selection criteria

It seems useful to give a checklist of selection criteria for the purchase or use of a 3D acquisition device. The different items follow those listed in a recent report on how to buy time-of-flight systems, published by Spar Point Ltd[22]

Work flow: Some products yield surfaces, others mere point clouds. In the former case the topological relations between 3D data points are available - although possibly erroneous at times - whereas in the latter they are not captured and have to be established for most visualisation purposes. Raw 3D data, be it surface representations or point clouds, are rarely the end goal. It therefore is important the check on the formats in which the data can be delivered and the software to carry out subsequent steps. What are the available tools for cleaning up, building complete models from patches (so-called `registration'), simplifying the data with minimal loss of quality (so-called `decimation'), etc.? Some point-cloud or surface editing software products work best with MicroStation, some with Intergraph tools or AVEVA, others with AutoCAD. How open is the data format? This is a concern if the post-processing software does not come with the hardware, or if one wants to keep flexibility.

Accuracy: What is the accuracy of the system? This will typically depend on the distance to be measured, as well as on the reflectance characteristics of the surface. This makes it difficult to make clear-cut statements about accuracy. Usually, providers specify an uncertainty interval in which the measurement will lie around the true value (`groundtruth') and this for multiple distances. There is an important qualitative difference between the accuracy for time-

[22] Spar Point Research LLC 85 Constitution Lane, Suite 2E, Danvers, MA 01923 USA 978-774-1102 We want to thank Lon Addison of Unesco for pointing out the activities of Spar Point in this area to us.

of-flight systems vs. triangulation based system, as mentioned in the previous section. These qualitative differences tend to render time-of-flight systems effective for large environments, and triangulation based methods for small and medium-scaled objects (cm-meter range).

Obviously, requiring more stringent accuracy specs tends to increase the price of the system. Cultural heritage stakeholders must therefore refrain from imposing needlessly high accuracy demands. Extra precision often comes as a cost, not only financially but also in terms of the completeness to which artefacts are scenes can be recorded. For instance, the hand-drawn sketches of pottery fragments are notoriously imprecise. Hence, it is not self-evident to suddenly require micron precision when going to digital 3D scanning solutions instead. In this respect, there also is a difference to be made between absolute and relative accuracy. The former is a kind of end-to-end precision and specifies the maximal error in the coordinates of any of the measured points with respect to an absolute world coordinate system, i.e. with respect to the same reference position. Whatever dimension or distance is measured, this precision should be met. Relative accuracy is less restrictive, but often the really relevant issue. How precise is the position of points with respect to those in their immediate surrounds. This precision will e.g. determine whether small details will be visible, like seals or other imprints on pottery, whether fractured parts can be digitally fitted, etc.

Dynamic range: What is the instrument's dynamic range - that is, the useful working range of the instrument in the conditions under which it will be operated? Working ranges strongly depend on the type of systems. The working range of time-of-flight 3D laser scanners starts around 1 meter, and can be as great as 1 kilometer. The working range also depends on the reflectivity of the surfaces that are being measured, on the intensity and wavelengths of the ambient lighting, and on humidity conditions. Typically dynamic range gets traded off against acquisition speed for these systems. Triangulation based, active scanners tend to have a much smaller working range. Some of these devices have a very specific working volume, and the distance the surface being measured typically is less than a couple of meters, usually much less. Active, triangulation based system also have difficulties beyond a couple of meters, as the projected patterns will be difficult to pick up by the camera. Passive, triangulation systems can be used at the largest range of distances, if the optics of the camera(s) can be adapted and if the camera positions are sufficiently far apart to get reasonable `disparities' (i.e. shifts of the points between images).

Speed of the device: This determines the time that is needed to collect the required data. Higher speeds may imply lower precision, e.g. as fewer measurements are averaged for the same point by a time-of-flight scanner. In the case of active, triangulation-based systems, there has been an important shift from scanning to short sequences of projected patterns, with one-shot operation as the limiting case. Needing fewer projections obviously creates the potential to be faster, but this is not necessarily the case, as one-shot systems will typically require more processing as systems using subsequent projections. Yet, the time it takes to gather the necessary input can also be an issue. If this time is very short, there will be less disturbing influences from objects moving in the scene.

Field of view / working volume: What field of view (FOV) or working volume does the instrument have? For time-of-flight devices, the FOV in different instruments can range from a window of 40 x 40 degrees, to 270 degrees vertical x 360 degrees horizontal, to almost (but never fully) 360 x 360. A larger FOV allows to scan a scene from fewer vantage points.

Several device have a programmable FOV, which allows to avoid unnecessary scanning. Active, triangulation based devices typically have a working volume that is quite specific. Hence, for different sizes of artefacts, different scanners may have to be used or larger objects may have to be captured as a mosaic of many, piecemeal scans. Passive techniques are quite flexible in this regard, as long as the baseline is not fixed (as with some stereo cameras or stereo rigs).

Operating environment: 3D scanning sometimes has to be performed under harsh conditions. As long as scanning has to be carried out within the highly controlled environment of a museum there may be no problem, but outdoor operation under adverse weather conditions (heat, humidity) may be impossible with certain types of devices. Typically, time-of-flight systems and cameras will pose fewer restrictions than the use of active, triangulation-based devices. The latter tend to be less rugged and may find it difficult to compete with the ambient sun light. It is also important to check the possibility to operate on the basis of batteries (instead of needing to plug in to the mains) and what the up-time and reloadability are. Very precise systems may even require temperature-controlled rooms.

Flexibility and ease-of-use: Portability of the devices is a definite advantage. This is a problem mainly with the larger active, triangulation based systems. Otherwise, most technologies have become increasingly portable over recent years. Shape-from-video or shape-from-stills techniques described earlier probably are the pinnacle in that sense. Another issue is calibration and setup time. If calibration has to be carried out repeatedly or even at the manufacturer's local or central outlier at times, or if the necessary procedures are time consuming, then this is a definite weakness.

Eye safety: Active apparatus may pose a danger in that the lighting may damage eye sight. This said, the allowable intensities and durations of exposure have not been specified in internationally accepted guidelines. In the U.S., the Center for Devices and Radiological Health (CDRH) of the Food and Drug Administration (FDA) have issued regulations. In Europe and elsewhere, regulations issued by the C.I.E. (Commission Internationale de l' Eclairage) apply.

This said, there are additional standards and directives, e.g. from the U.S. Department of Labor, Occupational Safety & Health Administration (OSHA), and the American National Standards Institute (ANSI). In any case, it is important to consider the risk of exposure - mainly to laser beams - of workers or the public. The risk is mainly present with time-of-flight devices as they tend to operate over larger distances, and are scanning around a central acquisition point, thereby covering a possibly large working volume. The risk is smaller but not necessarily non-existing with triangulation based, active scanners. On the other hand, their working volume tends to be smaller and better shielded off. Also, the intensity of the lighting is typically lower. On the other hand, they will more often operate in the visible range of the spectrum, which increases the sensitivity of the eye to the corresponding wavelengths (but renders the light visible and therefore at least a noticeable hazard). With passive 3D acquisition (purely from images taken under normal, ambient light) eye safety normally is no issue.

Sustainability: It is important that the supplier of the technology has a proven market record, and seems stable enough to be still around at the time when replacement parts or maintenance services are needed. Depending on the product, the costs of upgrading the system may also be very different. Time-of-flight scanners and active, triangulation based systems will typically need more drastic modifications of entire replacements, than passive, triangulation based systems, where the software will typically handle images of higher resolution right away, immediately yielding better models.

Cost: This obviously is a major concern. Time-of-flight devices typically are the most expensive. Prices start at about $30,000 and can be as much as $200,000 incl. software, maintenance, training and support. Active, triangulation based systems tend to be slightly cheaper, but not much. Prices start at about $20,000 but also go up a lot for dedicated body scanners (possibly useful for statuary, for instance).

7.1.1.3. Commercial suppliers

This section lists some system or service providers for 3D acquisition[23]. These lists are by no means complete, however.

Time-of-flight scanners: iQvolution/FARO, Leica Geosystems HDS (previously Cyrax), MDL, Optech, Riegl, Trimble, Visi Image, and Z+F

Active, triangulation based: 3D scanners, 3dMetrics, Cyberware, Eyetronics, Minolta, Polhemus, ShapeGrabber ...

Passive, triangulation based: 2d3, Eyetronics, PixelFarm, Realviz,... for software solutions based on video sequences, Point Grey and others for stereo cameras that yield depth maps.

7.1.2. Archaeological recording

The accurate recording of excavated features has been a problem for archaeologists since they have striven to define archaeology as a science. In recent years, the use of digital recording techniques has become increasingly widespread in archaeology. The rate of adoption of digital technologies in the field has been dependent upon both the level of resources available as the technical requirements. As the cost of equipment continues to drop the financial constraints on its use become less of an issue.

In most cases where digital recording is used, it is still used in parallel with conventional recording methods. Some aspects of the record have been captured "conventionally", and then digitised - so the conventional record still forms an intermediary in the recording process - while other aspects of the record may have been directly digitally captured with no intermediary. In addition, there are few projects which have tried to compile a complete digital record of the conventionally assembled record: for example, only a selection of the photographic images are subsequently scanned, or the plans are digitised but the section and context sheet sketches are not.

[23] SparView report `How to Buy a 3D Laser Scanner: Spar Point's Top 10 Checklist', Vol. 3, No.12, Tuesday April 5, 2005, ISSN 1553-8834

7.1.2.1. The Recording of Archaeological Features

Traditionally there has been, and to a large extent still is, a division between excavation and post-excavation activities. This boundary between these two has always been somewhat blurred and this is even more so with the use of computer-based recording and analysis which enables and encourages seamless integration between the two. Despite the pressures of time and resources that exist on most modern excavations the aim is to produce, on site, a descriptive written, drawn and photographic record for every context together with other linked records for artifacts, ecofacts, samples and other specialist materials. A part of this record is the Harris Matrix which, ideally, should be constructed and checked during excavation but in reality is often finished afterwards.

Alternatives to manual recording are becoming more popular and more widely-used, in the hopes of producing a system that overcomes the limitations of traditional data collection. While most of these systems solve the problems of accuracy and time, they tend to come with their own problems including expense, inconvenience and technical complexity.

The record of a site is thus traditionally subdivided into the written, drawn and photographic records. Computer-based recording and analysis is now enabling the integration of all three elements in new and exciting ways. Increasing digitization is causing a fundamental re-evaluation of what excavation archives are being created for with a shift from simple storage to an emphasis on access and re-use. Coordinated use of the Internet is an important element within these developments, often at a national scale with the development of digital archiving services such as the Archaeology Data Service (ADS) in the UK. The ADS Guide to Good Practice for Excavation explains in detail how to prepare a digital excavation archive to be accessible via the Internet (Richards and Robinson 2001).

7.1.2.2. The written record

The traditional approach to excavation recording, consisting of written descriptive passages, short notes and annotated sketches in day books and diaries, produces a record which is difficult to store and analyse on a computer. To accomplish this the information must be logical in both its structure and its intended meaning.

The Museum of London Archaeology Service (MOLAS) for example operates a very precisely defined computer-based recording system for its many excavations within London. This is well documented within the MOLAS recording manual (MOLAS 1994).

One of the major benefits of using computers is that it forces us to think logically about the way data are structured and the way we work with data. Excavation recording, however, has never been amenable to standardization and the introduction of computers has certainly not altered that.

7.1.2.3. The drawn record

The drawn record consists mostly of manually drawn records, and thus will not be discussed here, though even in this field there are some digital tools available. A German team for example developed a compact machine, called the *Kartomat,* capable of rendering scale drawings of both horizontal and vertical features in a short amount of time (Arnold and Gersbach 1995). Also, the utility of computer graphics by using CAD products for drawing, mapping, and 3-dimensional reconstruction in archaeological research has been widely recognized. Off-the-shelf image processing software and some CAD based software like have been used quite successfully for manipulating images, producing archaeological line drawings, and saving the drawings for further modification (Eiteljorg 1997).

7.1.2.4. The photographic record

The immediacy of digital imaging, the ease at which it can be used and subsequently reused, combines to make it a powerful new tool for the archaeologist. Now not only can the original excavation work be recorded using digital photography but also the transfer of this information can be made almost immediately. Archaeological projects nowadays invest a lot of effort to digitize their archaeological photographic record. The use of digital photography is not limited to still photography. One can make virtual tours of the site, using QuickTime VR. In addition often selected moments of the excavation progress are recorded using digital video. Although the use of digital imaging has great potential it does not come without a number of negative aspects, which have to be addressed to successfully incorporate this newest form of photography into the tools used by archaeologists (Mauzy 2003).

7.1.2.5. Harris Matrix

The physical relationships of a context with others around it in the ground are essential pieces of information to be recorded at the time of excavation. Once recorded, the whole stratigraphic sequence of the site can be represented diagrammatically in the form of a Harris Matrix. Such a matrix has become a fundamental tool in the interpretation of a site and is usually produced as a part of the site archive. Because, in essence, this is a simple and very logical concept often to be applied to large data-sets of many thousands of contexts, its potential for computerisation was soon realised and since the late 1970s several matrix programs have been developed with varying levels of success. Ryan (2001) gives an interesting overview of computerised stratigraphy.

The application of computers to the manipulation and analysis of stratigraphic sequences probably began with Wilcock's STRATA program (1975). STRATA demonstrated that a computer program could be used to derive a logical sequence from a collection of relationships between stratigraphic layers. The appearance of this program also led to the start of a debate about the appropriate use of computer based techniques in stratigraphic sequencing and, more generally, in excavation recording that was to continue for some time.

STRATA's input was a complete set of recorded observations of stratigraphic relationships, and its output was a complete sequence for the site. It was designed as a batch process in which all of the available data was interpreted in a single run to produce a deterministic solution. This 'black-box' approach was seen by many as antithetical to the process of developing an understanding of a site and its stratigraphy during excavation (see, for example, Harris 1975). In practice, contemporary hardware and software limitations severely restricted the size of the models that could be handled and processing even very small sites might take several hours. However, it was to be some time before interactive computer methods could offer an alternative approach.

Whilst some concentrated on the stratigraphic diagram, or 'Harris Matrix', as a distinct issue, others began to develop systems to integrate relationships and sequences into more complete excavation recording and analysis tools. Rains, for example, introduced the forerunner of his integrated archaeological database at this time (Rains 1985). Soon after, Alvey presented his Hindsite program which used AutoCAD to maintain single-context plans together with a representation of the relationships between layers (Alvey 1989). Both of these systems represented significant early contributions to the development of excavation recording and visualisation.

Herzog and Scollar (1991) introduced a fully automated system for producing stratigraphic diagrams. Though able to handle more realistic data volumes with reasonable speed, Herzog's program followed Wilcock's earlier approach of producing a solution as the output of a batch run. However, improving technology was later to make it realistic to run the program whenever new data became available during excavation. In this way, the sequence diagrams produced by the program could evolve as excavation progressed. The main limitation of this approach was that the excavator could exert little influence over the final form of the diagram.

Taking matrix generation one stage further is gnet which enables links between database records and contexts (Ryan 1995). The design thinking behind gnet was quite different to that behind Herzog's program. Whilst accepting that automated layout and the ability to print stratigraphic diagrams were important capabilities, the ability to interact with and explore the diagram were seen as the key to supporting the excavator's need to develop and maintain an intimate understanding of the structure of the site as excavation progressed. The philosophy behind gnet is important in illustrating the increasing trend towards integrated information systems utilising the inherent strengths of commercially available software. Gnet proposed a digital excavation archive with analytical capabilities centred on the site matrix utilising invisible software links to standard commercial packages such as, database, CAD and word-processing software. In the mean time a successor to gnet, called jnet, is under development at the University of Kent.

Beside the creation of a stratigraphic sequence (Harris matrix) the stratigraphic excavation method is based on the recording of every single unit of stratification. Conventionally it is done by analogue drawings. GIS can be used for the digital recording of stratigraphic excavations. During the last years GIS based procedures were developed for the digital documentation of stratigraphic excavations. To be able to fully reconstruct the part of the site destroyed by excavating, the surfaces of the excavated deposits have to be fully documented in 3D ("single surface mapping"). Boundary polygons, topography, location of finds, (rectified) pictures… can be imported in a Geographical Information System (e.g. ArcView GIS – ArcDig), see for example Doneus & Neubauer (2004).

7.1.2.6. Spatial data

One of the features of modern archaeological research is that many different kinds of data are relied on or collected during the process of excavation. The traditional divisions of excavation recording into written, drawn and photographic records however, are becoming increasingly blurred trough the use of IT. The recording of spatial coordinates, for example, once part of the written record, can now be inherent within spatial data put to a variety of uses.

Archaeological data are inherently spatial, and archaeologists are naturally concerned with the distribution of archaeological sites across the landscape. From these distributions, described as settlement patterns, we can infer a great deal about the social and political complexity of the ancient people's we study, the size of their domains, aspects of resource procurement, and much more.

Data collection or sources of data relied on to build information about a site increasingly are in digital form. These sources of data may include use of (Craig & Aldenderfer 2003):

- Space-borne remote sensing detectors for aiding in the establishment of a site within the larger landscape context.
- Geophysical survey instruments for subsurface remote sensing. These may include magnetometers and ground-penetrating radar instruments.
- Electronic total stations for mapping of both surface features and features exposed during excavation.

These methods of data collection improve our ability to rapidly and accurately record a deposit during the process of excavation. Space-borne remote sensing permits us to view a site and its surroundings from a perspective that is not obtainable from the ground. Geophysical instruments give us a view below the surface that allows us to guide the prospect of excavation. Total stations permit extremely accurate and rapid collection of spatial data, and digital cameras permit excavators to review photographs in the field so that one can be assured of the photographs' quality.

7.1.2.7. Underwater archaeology

Archaeological prospection and subsequent recording takes place beneath the sea as well as on the ground and from the air, with the search for submerged shipwrecks and other cultural remains. Regardless of whether the remains lie on the sea-bed or have been buried by sediment, a suite of sensors including sonar, side scan sonar, sub bottom profiler and magnetometer can be used from a ship systematically tracking across the area of survey (Blake 1991). All methods produce digital gridded data which can be manipulated in the same ways as other digital images. Underwater magnetometry is similar to that on land and is usually processed by looking at grey-scale images and wire-frame diagrams; sonar takes depth measurements directly beneath the ship which can be interpreted by contour maps or 3-D plots based on the locational coordinates of the ship at each reading.

Once structures are located underwater their recording presents problems not found on dry sites, not least the amount of time divers can spend working on them. This has resulted in the importance of using photomosaic techniques and photogrammetry for underwater recording.

In addition to physical and/or electronic measurements, a site survey includes the use of underwater photographic recording such as site photographs, a photomosaic, stereo-photogrammetry and/or video recording. Site survey methodology may also employ remote sensing equipment such as metal detectors, magnetometers and/or sub-bottom profilers.

Current underwater archaeological survey thus utilize a combination of the most recent applications in remote sensing technology and data processing capabilities. Nautical archaeology is undoubtedly on the forefront of the use of IT solutions in archaeology (e.g. Catsambis in press)

7.1.2.8. Some tools for archaeology

In order to conclude this section on archaeological recording, we have listed some tools developed especially for the purpose of assisting the professional archaeologists in recording the past. Some of these tools have already been introduced briefly in this section.

ArchEd[24] is a tool for drawing Harris matrices which are used in archaeology. Beside its ability to edit such drawings it also contains an automatic drawing feature which redraws a given graph nicely. A similar program was developed at the Amt für Bodendenkmalpflege in Bonn in 1990. While this older program runs on DOS - systems and uses the keyboard as input device, ArchEd runs with Win2000, XP or WinNT and can also use a mouse as input device.

Proleg Stratigraf[25] is the complete framework for field archaeologists that allows you to record, document, analyze and interpret archaeological sites. It integrates in a single place field registration data, photography, bibliography, artifacts inventory, the archaeological drawing and the Harris matrix, creating thus, a real Digital Model of the excavation. The model provides you with a vision of what the site is really like and therefore the ability to act and interpret based on totally reliable information. Various versions (LITE, PRO) provide progressively more powerful registry and interpretation tools.

Proleg MatrixBuilder[26] is a unique software for archaeologists that automatically generates a fully customizable and error-free Harris Matrix taking into account all context chronologies. It is a product ideally suited to those interested in adding an "off the shelf" matrix-drawing tool to their existing excavation recording system

The Bonn Archaeological Software Package (BASP)[27] is a non-profit software project for and by archaeologists which has been developed cooperatively since 1973. It now includes more than 70 functions for seriation, clustering, correspondance analysis, and mapping tools for archaeologists working with IBM compatible PC's under DOS and all versions of Windows. It

[24] http://aragorn.ads.tuwien.ac.at/ArchEd/

[25] http://www.proleg.com/stratigraf.asp

[26] http://www.proleg.com/mbuilder.asp

[27] http://www.uni-koeln.de/%7Eal001/basp.html

also includes programs for three dimensional display of data, for finding rectangular structures in scanned excavation plans containing thousands of postholes, and for the rectification of extremely oblique aerial photographs and their superimposition on large-scale scanned maps under Windows NT/2000 and Windows 95/98/ME.

gnet[28] - jnet[29] is a graph browser/editor that can be used for visualising, analysing and exploring archaeological stratigraphy and other directed graph-based data. Gnet development ceased several years ago. More recently an entirely new project began to develop jnet a program with similar functionality, but capable of working on a wide range of devices from handhelds to servers and supporting collaborative graph manipulation.

STRAT[30] has been designed for archaeological interpretation in 3D space. Strat has the capability of combining legacy stratigraphy data from notebooks with contemporary photogrammetric data.

ArcDig[31] is a Windows program designed to present information about an archaeological site in a new way. ArcDig is a program that allows you to dig down through a site's layers, discover and examine finds, and go to web pages giving further information about them.

SIDGEIPA[32] is mainly dedicated to the management and retrieval of the archaeological information, including a number of useful functionalities in order to assist the archaeologists as well as the fieldwork researchers in the tasks of storing, processing and exploiting the obtained information related to excavations. The tool is even now in continuous evolution, being improved with new functionalities as they are suggested by users.

ArchaeoPackPro![33] is a software package that unifies all the elements of computer usage during archeological campaigns and provide an interactive research tool for data analyses after campaign ends. It uses a unique graphics user interface in 3D environment, allowing simultaneous and interactive work with the 3D terrain model, Databases, Statistical and mathematical analysis, 3D models of archaeological finds, etc. It will open a new chapter in archaeological documentation and interactive fieldwork, offering an ultimate context preservation tool for versatile field documentation, its safekeeping, flexible presentation and publication possibilities for a new age archaeology.

[28] http://www.cs.kent.ac.uk/people/staff/nsr/arch/gnet/

[29] http://www.cs.kent.ac.uk/people/staff/nsr/arch/jnet/

[30] http://www.brunel.ac.uk/project/murale/strat.htm

[31] http://arcdig.sourceforge.net/

[32] http://www.uv.es/%7Eamapa/angles/viena.htm

[33] http://www.online-archaeology.com/ArchaeoPackProNew.htm

Re:discovery's Archaeology Module[34] provides a complete record-keeping and interpretive research database for artifacts and their original context (site, major site features, excavated opening, and level). The archaeological records are integrated with the objects database to manage cross-mended objects and a study collection. As with all Re:discovery modules, the screen format is flexible to best reflect the nature of the artifacts and the kind of information being tracked. Commercial software.

ArchaeoData[35] is an excavation database on Microsoft Access, developed since 1998 by ArcTron GmbH. This incorporates project management and finds processing, a scientifically defined finds book, and image processing, including the systematic recording of archaeological data. This database processes a wide ranging level of reports, as well as handling plans and profile sections. The description formulas are developed around various archaeological criteria, such as pedology. The database can search and print out detailed archive material quickly via different adjustable data sourcing. Commercial software.

34 http://www.rediscov.com/archaeology.htm

35 http://www.arctron.com/Software/ArchaeoDATA/index.php

7.1.3. Archaeogeophysical prospection

7.1.3.1. Introduction

Decades of geophysical investigations for archaeological purposes have produced a series of useful information for academic archaeological analyses as well as for the protection of the archaeological cultural heritage from construction interventions. In the last decade geophysical prospection has become an inseparable part of the comprehensive archaeological research. In connection with the excavation works or probing they secure the optimum solution of the required tasks within the systematic research and rescue operations from the points of view of technique, time and economy. The choice of the optimum complex of geophysical methods is generally rather complicated and it cannot always be stated exactly for every task to be solved.

Geophysical techniques are **non-destructive**, since no digging or drilling is involved and all the necessary information is obtained above the ground. Especially the non-destructive aspect of geophysics turned out to be very promising, enabling archaeologists to investigate buried remains without damaging them. Thanks to geophysics archaeological sites can be localized both quicker and cheaper. Subsequently an inventory of archaeological sites can be made, enabling authorities to protect them against future destruction by building activities.

The vast majority of professional geophysicists work in areas such as geology, mineral prospection and hazardous waste management. In general, these subfields are concerned with the detection of large-scale and/or high-contrast targets. Thus, geophysicists trained in non-archaeological disciplines are generally not well-qualified to design surveys capable of detecting numerous types of archaeological features.

The term archaeogeophysical survey is employed here to indicate those techniques, that use measurements of physical fields at the surface of the Earth to investigate shallow depths, normally down to a few meters. Most geophysical techniques used today for archaeological investigations were initially developed for geological or civil engineering applications. While the underlying physical principles are the same, the shallow depth and relatively small size of archaeological features test the laws of geophysics to limits that are not often encountered by “conventional” geophysicists. Research into these particular problems led to the development of a sub-discipline now called “archaeological geophysics”. Doesn’t matter how correct that statement is, shouldn’t take us away from the fact, that we are still dealing with geophysical method and at least some procedures developed and frequently applied in the field of “conventional” geophysics have to be taken into account during designing of archaeogeophysical strategy. The most obvious approaches which should considered are: **Geophysical modelling** (Desvignes et al. 1999; Tsokaset al. 2000; Coskun et al. 2000; Eppelbaum et al. 2001; Hašek 1999) and **Simulating of anomalies** in different physical fields (Zdanovich et al. 2003; De la Vega et al. 1995), introducing of **Inversion metods** (Diamanti et al. 2005; Desvignes et al. 1999; Hašek 1999), **Deconvolution methods** (Karousova 1979), **Quantitative integration of geophysical methods** (Piro et al. 2000) at the interpretative level, **Composite images** are interpreted by statistical unsupervised classification approach (Ladefoged et al. 1995), **Neural networks** are used to derive estimates of feature burial depths (Bescoby et al. 1994), **On Pole and Pseudogravity transformations** to determine actual position of targets, **Time slices** are introduced in georadar data presentation (Nishimura 2000; Neubauer et al. 2002; Linford 2004).

Evidence of increase in the use of geophysical techniques by archaeologists are perceived in an increased number of World Wide Web sites that illustrate and summarize the results of a wide range of geophysical surveys. The most completed list with resources relevant to Archaeological prospection is a part of service maintained by the Department of Archaeological Sciences, University of Bradford[36].

A number of recent monographs and articles also document the increased use of geophysics (Clark 1996, Conyers and Goodman 1997, Bevan 1998, Hašek 1999; Clay 2000, Silliman et al. 2000, Kvamme 2001, Hargrave et al. 2002, Gaffney et al. 2003).

Techniques of Detection in Archaeological Geophysics

Method	**Active or Passive**	**Frequency of Use**
Electrical resistance	Active	High
Magnetometry	Passive	High
Electromagnetic	Active	Mid/Low
Magnetic susceptibility	Active	Mid/Low
Metal detektors	Active	Low
Ground Penetrating Radar	Active	High/Mid
Seismic	Active	Low
Microgravity	Passive	Low
Induced polarisation	Active	Low
Self Potential	Passive	Low

[36] http://www.brad.ac.uk/acad/archsci/subject/archpros/archp_nf.php

7.1.3.2. Commercial geophysical processing software

Geoplot 3.0 (Geoscan Research) – Professional archaeogeophysical survey software[37].

Geoplot 3.0 is a Windows program for the processing and presentation of geophysical data collected from a variety of instruments including: resistance meters, gradiometers, magnetometers, EM instruments and magnetic susceptibility instruments. Processing facilities include: high pass, low pass, median & periodic filters, spectrum and variance analysis, despiking, interpolation, edge matching, zero mean traverse correction, destagger correction, several numeric functions and a powerful cut and combine function for combing data sets mathematically. Graphics may be produced as shade plots (grey scale or colour), trace plots (stacked profiles or 3D), dot-density or pattern plots and may be printed out at any scale or saved as bitmaps for use in other software packages. A publishing mode is included which allows you to combine many graphics images, text, drawn objects etc. Data can be imported and exported, allowing data exchange with other software packages (Surfer, Golden Software Inc.; Idrisi, Clark University).

InSite – (Geoquest Associates) – Professional archaeogeophysical survey software.
The InSite program has been developed to enable the downloading, processing and visualization of achaeogeophysical and topographic data. InSite uses a layer model to manage and display site features, grid maps and images together with their geophysical and archaeological interpretations. Up to 3 geophysical or topographic data sets can be manipulated simultaneously and be combined using logical bitmap or arithmetic operators to explore reduction of artifacts specific to magnetometer data such as striping and staggering on zigzag traverses, instrument drift and periodic defects due to instrument oscillations or operator magnetization.

Archeo Surveyor 1.2.5.0 (DW Consulting B.V.) – Professional archaeogeophysical magnetometer and resistivity software. 30 days free TRIAL period version with full capabilities available from the web site[38]:

This software is specifically designed for the needs of archaeological prospection. ArchaeoSurveyor is a complete package for downloading, assembling, enhancing and saving the geophysical data from a range of instruments (Geoscan Research, Bartington and TR/CIA). Many views of the data are instantly available including: Block & Graduated Shade, Trace, 3D, Spectrum and Spredsheet, Shade and Trace views can be rotated to any angle.

GM-SYS Profile (NGA) – Software for magnetic modelling. DEMO version offered[39].

This is an interactive geophysical magnetic modelling program that enables us to rapidly create a geologic model and test its accuracy by comparing the model's magnetic response to observed gravity and magnetic measurements. For magnetic modelling of archaeological objects which are rather small and inexpensive, Lab License meet majority needs. Advantage for archaeogeophysical modelling is that it allows us to import bitmap images of georadar profiles, what enables us to create more reliable magnetic models of archaeological objects depicted on the radargrams.

---

[37] http://www.geoscan-research.co.uk/page9.html

[38] http://www.dwconsulting.nl/archeosurveyor.htm

[39] http://www.nga.com/GM-SYS-Profile.htm

Oasis montaj (NGA) – Powerful mapping and processing software for exploration and earth science investigations including environmental projects. On line DEMO[40].

Variety of Montay extensions and Montay plus extensions are available for advanced geophysics and geochemical data processing, analysis, quality control and modelling, making this software one of the most robust and comprehensive geoscience mapping and processing system in the world. Its present acceptance in archaeogeophysical prospecting is zero. Anyway, its advantages should be examined seriously at least for planning of archaeological prospections on the huge and complex archaeological sites, like ancient cities are.

MOD3D (Igor Cerovsky) - Interactive 3D geophysical gravity and magnetic modelling software. DEMO version offered[41].

Mod3D has been developed to create 3D geophysical models in an user friendly interactive environment. Anomalous gravity and/or magnetic fields are computed using formulae for polyhedral bodies. Geo-referenced bitmaps of georadar profiles or geoelectric pseudosections could be imported, and drawn transparently. This program is still under construction but its seams to be the best choice for archaeogeophysical 3D magnetic modelling.

Magix Plus v3 (Interpex Ltd.) – Software for magnetic modelling[42].

Software tool that interpret magnetic profile data. It is based on data models in 2, 2.25 and 2.75 dimensional earth models consisting of bodies and layers whose forward response matches the input data along a single profile line. The program use a generalized "2.75 dimensional" model to compute the theoretical magnetic response. The 2.75 D model represents the subsurface features as polygonal prisms with horizontal axes and either finite or infinite extent in the strike direction.

Radan 6.0 for Windows (Geophysical Survey Systems, Inc.) – Professional GPR post-survey processing software[43].

New Radan represents the most complete set of required facilities for radar data processing on the market. Recently added and unique feature among the georadar software is feasibility of collecting and integrating GPS data with the corresponding radar files. GPS data is stored in convenient Microsoft Access format. New *Map mode* allows 2D files collected with GPS data to be viewed in 3D and automatically estimate maximum depth capability.

- **QuickDraw Super 3D module**: The add-on module features 3-D presentation of data with simple manipulations of the entire data "cube" so that it can be "diced and sliced" along various x-y, y-z or x-z planes.

Reflex-Win (Sandmeier Software) - Professional GPR post-survey processing software.
Reflex-Win is program for the processing and interpretation of reflection and transmission data (special applications: ground penetrating radar (GPR), reflection seismics, refraction seismics and

[40] http://www.geosoft.com/pinfo/oasismontaj/demo.asp

[41] http://homepages.uni-tuebingen.de/igor.cerovsky/Main_Soft.html

[42] http://www.interpex.com/dos_programs/mag.htm

[43] http://www.geophysical.com/software.htm

ultrasound). Beside the Radan (GSSI) and Sensor's and Sofware GPR data processing modules, this is the most powerful program for GPR data processing and visualization. The most appreciated facility for archaeological prospection is user friendly, fast and efficient creating of horizontal cross sections (time slices). The slices are either displayed in manually scalable windows or by moving through the 3D-cube using a track bar.

Sensors & Software GPR data processing modules – Professional GPR post-survey processing software[44].

EKKO_View module – Provides powerful data plotting capability.
EKKO_View Enhanced and EKKO_View Deluxe – This module is designed for the GPR user who needs to edit, process, and plot large volumes of GPR data quickly and efficiently.
EKKO_3D – Provides 3D visualization of GPR grid data. Fascinating 3D slice, cube, and chair presentations and movies are created quickly and easily to aid in understanding complex three dimensional structures.

Slicer Dicer (Pixotec LLC) - Volumetric data visualization software. 14 days TRIAL version available[45].

The main power of this program is visual exploration of multidimensional volume data by "slicing and dicing" to create arbitrary orthogonal and oblique slices, rectilinear blocks and cutouts, isosurfaces, and projected volumes. Generate animation sequences featuring continuous rotation, moving slices, blocks, parametric variation (time animation), oblique slice rotation, and varying transparency. Significant advantage for archaeological geophysics is in visualizing sets of georadar parallel profiles[46] and geoelectrical pseudosections [47].

WinFence V2 (GAEA Technologies Ltd.) – Program for creating fence diagrams. 14 days TRIAL version available[48].

WinFence can be used to graphically create detailed, full colour, cross-sections and fence diagrams quickly and esily. The program can be used to interpret and map georadar profiles and geoelectrical pseudosections. Aside of this, it's offers possibility to link together soil horizons and/or archaeological stratigraphic units ascertained from boreholes which come after geophysical survey.

Idrisi Kilimanjaro (Clark University) – Low cost GIS and image processing software[49].

Idrisi is one of the most widely distributed low cost raster GIS and Image Processing systems in the world. It provides the most extensive set of GIS and Image Processing tools available in a single and

---

[44] http://www.sensoft.ca/products/index.html

[45] http://www.slicerdicer.com/

[46] http://www.lb.netic.de/hvdosten/

[47] http://www.geometrics.com/geoelectrical/geoelectrical.html

[48] http://www.gaea.ca/en-us/dept_2.html

[49] http://www.clarklabs.org/Home.asp

affordable integrated package. Idrisi covering the full spectrum of GIS and Remote Sensing needs from database query, to spatial modelling, to image enhancement and classification. For Image Processing, a complete suite of tools is available for restoration, enhancement and transformation, and for signature development and classification, including hard and soft classifiers and hyperspectral image classification. Statistical module contains important GIS and remote sensing algorithms. In archaeological geophysics, unsupervised classification for instance can be used at the interpretative level. Together with the composite module, which combines readings from different geophysical instruments, into a single file, is used to build up a site map in a single image, where archaeologically relevant units are discernable by spatially defined classes, based on the level of correlation between input data sets. The output is a composite image of differing, but normalised input data, and that could lead to partial automation of the interpretation procedure. The results will vary in terms of quality, but they enable a comparison of the individual data layers and the generation of new information based upon the level of correlation between them.

ER Mapper 6.0 (ER Mapper Pty Ltd.) – GIS and image processing software with built in geophysical data imaging and presentation module. DEMO version available[50].

ER Mapper has the ability to process geophysical data, in combination with other software, and enhance geological interpretation through its unique image processing capabilities. The imaging can be an integral part of the interpretation process. Throughout, the over-riding theme is to use geophysical images, preferably using all available data in an integrated multiband dataset and with vector overlays showing the original data distribution (bias) from which the image was created by gridding. The latter is important to enable recognition of meaningful data and spatial frequencies (as opposed to gridding artefacts). The use of an empirical and creative approach to combining and interpreting geophysical data is made possible by the power of ER Mapper.

EVS (C Tech Development Corporation) - Environmental visualisation system. DEMO version available[51].

C Tech unites state-of-the art analysis and visualization tools into extremely powerful software systems developed to meet the needs of geologists, geochemists, environmental or mining engineers, oceanographers, archaeologists and modellers. It provides true 3D volumetric modelling, analysis and visualization. Environmental issues can quickly become extremely complex. Maintaining and organizing that data is insufficient. Visualization is the only means for condensing and communicating vast quantities of spatially referenced data.

7.1.3.3. Freeware geophysical processing software

Snuffler 0.6 (David Staveley) - Freeware geophysical software designed for archaeological Resistivity and Magnetometry[52].

Enables data import from Geoscan RM15 and TRSystems meters and all indispensable and important steps in data processing (despike, interpolation, high pass filters - geological flattening, sharpen, edge

[50] http://www.ermapper.com/

[51] http://www.ctech.com/

[52] http://www.homeusers.prestel.co.uk/aspen/sussex/snuffler.html

correction and destripe). Image view module supports all three frequently applied data presentation methods in archaeological geophysics: grey scale, dot density and relief plots.

TR Resistance Meter Interface Program V1.33 - Freeware geophysical software designed for TR/CIA resistance meter[53].

PC software for Microsoft Windows 95/98/ME/NT/XP has been provided for downloading, displaying and editing of data for (low cost) TR/CIA Resistance Meter[54]. The software allows multiple grids to be merged and displayed with the images being stored as a standard windows bitmap file (*.bmp). Further data processing can be completed using normal image processing software such as Adobe Elements (not supplied).

MagMap2000 4.56 (Geometrics Ltd.) – Freeware geophysical software for magnetic method. It is simple-to-use interface program for loading data, making GPS, profile and contour plot[55].

MagMap is representing tool that reads GPS/Mag records and converts them into UTM/Transverse Mercator system based on standard or custom ellipsoids. Exports results in the format suitable for other gridding and contouring programs (Surfer of Golden Software Inc.) and for usage by freeware Magpick (Geometrics Ltd.) software for magnetic data processing and interpretation. It is post acquisition processing software used to analyze and process data from a number of different instruments. It allows the user to download data from the G-858, and modify the positions that were entered while doing a survey. This operation is not included in any other freeware or commercial software developed for archaeological magnetic prospection. The editing capabilities of these software should be exposed as the particularly applicable routines for repositioning of field readings.

MagPick V 2.6 (Geometrics Ltd.) – Freeware geophysical software for magnetic processing and interpretation. Sophisticated integrated data processing program with map overlay, contour, filtering, target analysis[56]

This free powerful data processing and modelling package provides data base management, file importation, gridding, contouring and dipole pattern matching source body modelling for the experienced geophysicist or potential field data processor. Output includes a table of target anomaly locations and depths. Includes pre- processing of GPS locations and data reformatting. It provides visual analysis of the magnetic map and manual or automatic selection of anomalies. Includes: Viewing the magnetic map with dynamic adjustment of colour scale in accordance with min/max of the data. Program allows enlarging or zooming into different parts of the map and presenting zoomed sections in separate windows with automatic change of the colour scale to fit the data range. It does also a quantitative interpretation based on fitting of the observed magnetic anomaly using the field of a simple source. Both map or profile information can be used, and geometrical and magnetic parameters of the source can be estimated.

[53] http://www.trsystem.demon.co.uk/html/support.html

[54] http://www.trsystem.demon.co.uk/html/resistance_meter.html

[55] http://www.geometrics.com/Downloads/downloads.html

[56] http://www.geometrics.com/Downloads/downloads.html

GPR_process 2.2 (Larry Conyers, University of Denver) – Freeware georadar software designed for creating time slices[57].

GPR data processing software and example radar profiles from Petra (Jordan) by L.Conyers. The ***GPR_Process*** program, which produces amplitude slice-maps, was written with archaeologists in mind. It is really a series of windows menus, overlaying a more sophisticated C++ program that was written to run only in DOS. At the present time this program is written only for users of GSSI (Geophysical Survey Systems, Inc.) data ***not post-processed with RADAN***. It is also important to understand that GPR data from Sensors and Software and RAMAC systems can be easily converted to GSSI format (which is .dzt format data), with a program that Jeff Lucius wrote. The files from this program are nothing more than x,y,z data for slices in the ground, in ASCII format. These can easily be imported into any spreadsheet program for manipulation and finally ***Surfer*** (Golden Software Inc.), or other mapping programs, for sophisticated gridding and image production.

Groradar ™ 2003,11 (Garry R. Olhoeft) – Freeware georadar software for DOS with modelling features[58].

Freeware software for acquisition, processing, modelling and display of ground penetrating radar data by Gary R. Olhoeft for GSSI (*.dzt), Sensors & Software (*.dt1) and Mala RAMAC (*.rd3) data formats. This program is designed to acquire, process, model and display ground penetrating radar data. Processing includes a variety of filters and transforms to improve the data or remove defects, the ability to model finite size scatters to estimate velocity and depth, and a scroll able display to show the context when selecting a wavelet scan for full waveform modelling. Full waveform modelling allows the determination of subsurface electrical and magnetic properties, and demonstrates how these properties alter the signature of a radar scan.

RadView (Geophysical Survey Systems, Inc.) – The Radan files viewer[59]

RadView allows you just to open and view Radan files. It doesn't offer you to perform any other powerful processing function.

Rtoaw (Geophysical Survey Systems, Inc.) – The Radan to Ascii converson utility[60]

Rtoaw is a utility that allows data to be converted from Radan *.dzt format to Tab Delimited Ascii text with data stored in either rows or columns, either signed or unsigned values.

Rad2Bmp (Geophysical Survey Systems, Inc.) – The Radan to Bitmap conversion utility[61]

Rad2Bmp is a conversion utility allows the users to convert 8 bit and 16 bit Radan (*.dzt) files into 24 bit Windows colour bitmaps.

---

[57] http://www.du.edu/%7Elconyers/gpr_data_processing_example.htm

[58] http://www.g-p-r.com/

[59] http://www.geophysical.com/softwareutilities.htm

[60] http://www.geophysical.com/softwareutilities.htm

[61] http://www.geophysical.com/softwareutilities.htm

Geosect, Geoplan and Multigeoplan (Dimitrij Mlekuž) – Freeware programs for creating georadar time slices[62].

- **Geosect.pl**: Converts single georadar profile in ascii data into Golden Software's Surfer *.dat file format.
- **Geoplan.pl**: Excecutes single horizontal cross-section at the specified "depth" through the set of parallel georadar profiles in ascii data and save the result in Golden Software's Surfer *.dat file format. In this way georadar "time slice" at the specified "depth" is created.
- **Multigeoplan.pl**: Executes series of horizontal cross-sections through the set of parallel georadar profiles in ascii data at the several successive "depths" specified by the starting "depth" and selected step. Result is series of georadar "time slices" from top to bottom of the profiles in specified interval.

Para View (Kitware Inc) – Freeware Visualization Application[63].

ParaView is an application designed with the need to visualize large data sets in mind. The goals of the ParaView project include the following:
- Develop an open-source, multi-platform visualization application.
- Support distributed computation models to process large data sets.
- Create an open, flexible, and intuitive user interface.
- Develop an extensible architecture based on open standards.

7.1.3.4. Freeware tools supporting archaeogeophysical survey

ATAGS 1.0 (US Army Corps of Engineers) – Automated Tool for Archaeogeophysical Survey[64].

Description of the program is in report: Geophysical survey in Archaeology: Guidence for Surveyors and Sponsors (by Lewis E. Somers and Michael L. Hargrave). The ATAGS software tool allows the user to develop an effective survey design for a geophysical survey at a particular site. ATAGS produces a detailed report that also provides guidance on project management. ATAGS is presently designed for use in the Midwest and Plains regions of the United States. The survey designs are intended for geophysical instruments manufactured by Geoscan Research, but can also be used (with minor revision) with all comparable instruments.

GMT (The Generic Mapping by Tools Paul Wessel and Walter H. F. and National Science Foundation)[65]

[62] http://www2.arnes.si/~dmleku2/geoph.html

[63] http://www.paraview.org/HTML/Index.html

[64] https://www.denix.osd.mil/denix/Public/Library/NCR/archaeology.html

[65] http://gmt.soest.hawaii.edu/

GMT is an open source collection of ~60 tools for manipulating geographic and Cartesian data sets (including filtering, trend fitting, gridding, projecting, etc.) and producing Encapsulated PostScript File (EPS) illustrations ranging from simple x-y plots via contour maps to artificially illuminated surfaces and 3-D perspective views. GMT supports ~30 map projections and transformations and comes with support data such as coastlines, rivers, and political boundaries. For more detail information about GMT see: Wessel, P. and W.H.F. Smith 1991. Free software helps map and display data. EOS Trans. AGU, 72, 441.

Mortindale's Calculators On-Line Center – On-Line Earth's Magnetic Field Calculators for every day Magnetic Prospections[66].

Magnetic field calculators, applets, animations and simulations.

7.1.3.5. Freeware geophysical educational software

Charge (Garry R. Olhoeft) - Free educational software for DOS[67].

External sinewave electrical field stimulus applied to two opposite sign charges moving with finite velocity as a function of frequency of stimulus. Animation shows amplitude and phase of charge separation (polarization) motion fit single relaxation equation. Discusses consequences of frequency dependence in terms of dispersion in wave propagation (demonstrated in GRORADAR™ program[68]).

Geoelectric educational software. Simulation of resistance anomalies for Twin probes configuration. An example of Macromedia Authorware 3.5 Shockwave applications by Paul Cheetham[69] (Department of archaeological science, University of Bradford).[70]

Geophysical modelling or simulation of resistance anomalies in this specific case is frequently applied for determining depth, shape and preservation level of archaeological architectural remains. This approach can be also used at the initial level of preparing field strategy design. Simulation of resistance anomalies suggests the proper mobile probe separation for the specific conditions on the site with the regard to type of archaeological remains, their depth, dimensions and resistivity contrast to surrounding soils.

7.1.3.6. The GIS databases for geophysical prospecting

Due to the enormous amounts of data, they should be linked in such a manner that immediate access to information, essential for either academic settlement analyses or for the protection and conservation of the archaeological cultural heritage, is enabled. Much attention has been dedicated to national archaeological databases, in the past few years, that also incorporate

[66] http://www.martindalecenter.com/Calculators1B_7_Nav.html#NAV-MAG

[67] http://www.g-p-r.com/educatio.htm

[68] http://www.g-p-r.com/

[69] p.n.cheetham@bradford.ac.uk

[70] http://www.gla.ac.uk/Inter/Computerpast/archtltp/geosurvey/geosurv.html

results from geophysical survey. This could be an indication that archaeological prospecting, and consequently also geophysical survey, have sufficiently asserted themeselves throughout the academic world serving for archaeological settlement analyses, as well as for the everyday protection of the cultural heritage.

Newer databases are created in program packages that serve as tools for examining extensive databases for geographic information systems (i.e. ArcView, Esri; EVS). The primary adventage of these databases is that they enable a link between the textual part of the database and the graphic foundations composed of georeferential information on geophysical investigations and aerial photography (see i.e. Doneus and Neubauer 1998, 29-56). While establishing database for geophysical prospecting, the fundamental principles of the English Heritage Geophysical Survey Database (SDB), proposed by the Ancient Monuments Laboratory in 1994, have to be taken into consideration[71].

[71] http://www.eng-h.gov.uk/SDB/AboutSDB.htm/

7.2. Future directions and EPOCH's contribution

3D scanning technologies have evolved considerably over the last years, as already highlighted before. Yet, these developments involved mainly non CH areas: reverse engineering and entertainment. CH oriented 3D scanning comes with a number of requirements on its own, however. Within the same excavation campaign or museum collection, objects of very different sizes may have to be scanned. Also not only the shape but also the appearance of objects is crucial and something one has to capture with great detail. CH also has to deal with the most complicated and intricate types of surfaces and materials (jewelry, fur, textiles, etc.). In contrast to what seems to be assumed at times, 3D scanning is far from a closed chapter, especially for CH. The confusion derives from the same tendency to simply copy solutions from other domains and as they are satisfactory there, assume that therefore the technology is sufficiently mature for CH as well. Here is a list of developments that still need to be developed into well streamlined products and processing pipelines in order to effectively cater for CH 3D scanning needs.
On-line scanning: 3D models should be built up during the scanning, The model should be shown on a screen while performing the scanning, as it grows. One can then immediately inspect the completeness and the quality of the result. Now, the quality of the model can only be assessed after post-processing, while during scanning at most separate 3D patches are visualised. Off.-line data processing has the disadvantage that the lack of certain data (i.e. the presence of holes in the model) or erroneous data only become clear when it is too late. Having to return to the site were the scanning has to be done can be very time and money consuming.

Integrated shape and surface reflectance modelling: 3D scanners yield 3D shapes and sometimes also surface textures. But simply mapping the latter onto the geometry is still a far cry from extracting the precise radiometric properties of the surfaces, which are nevertheless crucial to produce renderings of the true appearance of the object. Currently, there are devices that yield detailed geometry or detailed surface reflectance characteristics. What would be needed is an integrated approach, where one process yields to the simultaneous capture of both. As a matter of fact, this also makes a lot of sense from a technical point of view, as knowledge about one yields strong cues about the other. This said, additional complications like transparency and translucency offer challenges that are very hard to solve.

Opportunistic 3D modelling: currently there is no single 3D acquisition technique that can deal with all types of shapes and surfaces. Methods may have difficulties with narrow structures, deep cavities, specular reflections, textured surface, untextured surfaces, etc. Sometimes, objects or scenes are adapted to fit the requirements of the 3D scanning device, e.g. by painting or powdering them. This is all the more unacceptable in a cultural heritage setting. Future systems should be hybrids that deal with much broader classes of objects by (automatically) adapting their modelling strategy to the object or scene.

Automatic registrations: typically, 3D data come in as partial patches. These then have to be glued together to form complete models. There are good solutions for fine-registration, but these require good initialisations, which are usually obtained by first manually puzzling the pieces together in more or less correct relative positions. Fine docking is then carried out automatically. This crude puzzling also ought to be done automatically in order to automate the entire process, i.e. there still is a need for robust, crude registration algorithms.

7.3. References

Alvey, B., 1989, ‘Hindsite’, *Archaeological Computing Newsletter*, 19: 4–5.

Ancona, M., G. Dodero, M. Mongiardino and A. Traverso, 1999, ‘Taking Digital Notes in the Field’, in J.A. Barceló, I. Briz and A. Vila (eds.), *New Techniques for Old Times, Computer Applications and Quantitative Methods in Archaeology, 1998*, pp117–121, BAR International Series, 757, Archaeopress, Oxford, UK.

Arnold, B. and E. Gersbach. 1995 The Kartomat: A field drawing machine. *Journal of Field Archaeology* 22(3):369-376.

Bescoby, D. J., Gavin C. C. and P. N. Chroston 2004. Enhanced interpretation of magnetic survey data using artificial neural networks: a case study from Butrint, southern Albania. Archaeological Prospection 11 (4), 189-199.

Bevan, B. W. 1998. Geophysical exploration for archaeology: An introduction to geophysical exploration. Midwest archaeological center Special report No. 1, U.S. Department of the interior, National park service, Linkoln, NE.

Blake, V.S. 1991. Remote sensing in underwater archaeology. *Science and Archaeology* 33: 3-17.

Catsambis, A. et al. (in press). Underwater surveys and remote sensing – the 2005 Corinthian Gulf Expedition. *Proceedings of CAA 2005*, Tomar.

Clark, A. 1996, Seeing Beneath the Soil: B.T. Batsford Ltd, London, 192 pp.

Clay, R. B. (http://www.crai-ky.com/geophysical/wheredow.htm) 2000. The place of geophysical survey in contemporary fieldwork. WWW page mainained by cultural resource analyst, Inc., 143 Walton Ave, Lexington, KY.

Conyers, L. B. and D. Goodman 1997. Ground-penetrating radar. An introduction for archaeologists. Altamira press, Walnut Creek, CA.

Coskun, N. and J. E. Szymanski 2000. A discussion on the resolution of two-dimensional resistivity modeling. Archaeological Prospection 6 (4), 179-186.

Craig, N. and M. Aldenderfer. 2003. Preliminary Stages in the Development of a Real-Time Digital Data Recording System for Archaeological Excavation Using ArcView GIS 3.1. *Journal of GIS in Archaeology*, volume 1: 13-22.

De la Vega, M., Osella, A., Lascano, E. and J. M. Carcione 2005. Ground-penetrating radar and geo-electrical simulations of data from the Floridablanca archaeological site. Archaeological Prospection 12 (1), 19-30.

Desvignes, G., Tabbagh, A. and C. Benech 1999. The determination of the depth of magnetic anomaly sources. Archaeological Prospection 6 (2), 85-105.

Diamanti, N. G., Tsokas, G. N., Tsourlos, P. I. and A. Vafidis 2005. Integrated interpretation of geophysical data in the archaeological site of Europos (northern Greece). Archaeological Prospection 12 (2), 79-91.

Doneus, M. and W. Neubauer 1998. 2D combination of prospection data. Archaeological prospection 5 (1), 29-56.

Doneus, M., and W. Neubauer. 2004. Digital recording of stratigraphic excavations. in *CAA2003: Computer Applications and Quantitative Methods in Archaeology*: 113-116

Eiteljorg, H. 1997. A New Aid for Drawing Pottery Profiles. *CSA Newsletter* IX, 4: . http://www.csanet.org/newsletter/feb97/nl029708.html

Eppelbaum, L. V., Khesin, B. E. and S. E. Itkis 2001. Prompt magnetic investigations of archaeological remains in areas of infrastructure development: Israeli Experience. Archaeological Prospection 8 (2), 163-185.

Gaffney, C. & J. Gater 2003. Revealing the buried past. Geophysics for archaeologists. Tempus Publishing Ltd, pp. 192.

Haigh, J.G.B. 1985, 'The Harris Matrix as a Partially Ordered Set', in E. Webb (ed.) *Computer Applications in Archaeology, 1985*, pp 81–90, University of London Institute of Archaeology, London, UK.

Hargrave, M. L., Domers, L. E., Larson, T. K., Shields, R. and J. Dendy 2002. The role of resistivity survey in historic site assessment and management: an example from Fort Riley, Kansas. Historical archaeology 36 (4), 89-110.

Harris, E.C. 1975, 'Stratigraphic Analysis and the Computer', *Computer Applications in Archaeology, 1975*, pp33–40.

Hašek, V. 1999. Methodology of geophysical research in archaeology. Bar International Series 769, 127 pp.

Herzog, I., and I. Scollar 1991, 'A New Graph Theoretic Oriented Program for Harris Matrix Analysis', in K. Lockyear and S.P.Q. Rahtz (eds.) *Computer Applications and Quantitative Methods in Archaeology, 1990*, pp53–59, British Archaeological Reports, Oxford, UK.

Karousova, O. 1979. Dekonvoluce profilovych krivek T (dipl. prace). MS PrF UK Praha.

Kvamme, K.L. 2001, Current Practices in Archaeogeophysics: Magnetcs, Resistivity, Conductivity, and Ground-Penetrating Radar, in Goldberg, P., Holliday, V and Ferrng, R. (eds.), Earth Sciences and Archaeology: Kluwer/PlenumPublishers, New York, pp 353-384.

Linford, N. 2004. From hypocaust to hyperbola: Ground-penetrating radar surveys over mainly Roman remains in the UK. Archaeological Prospection 11 (3), 237-246.

Lock, G.. 2003. Using computers in archaeology. Towards Virtual Pasts.

Mauzy, C. 2003. The use of digital photography in the field and studio-benefits and drawbacks. In *CAA2002: Computer Applications and Quantitative Methods in Archaeology*:

Mušič, B. and B. Slapšak 1998. GIS in on-site analysis: Rodik, Slovenia. – In: The use of geographic information systems in the study of ancient landscapes and features related to ancient land use (ur.: Peterson, J.). Cost action G2; Paysages antiques et structures rurales, 81-93.

Neubauer, W., Eder-Hinterleitner, A., Seren, S. and P. Melichar 2002. Georadar in the Roman Civil Town Carnuntum, Austria: An Approach for Archaeological Interpretation of GPR Data. Archaeological Prospection 9, 135-156.

Nishimura, Y. and D. Goodman 2000. Ground-penetrating Radar Survey at Wroxeter. Archaeological prospection 2, 101-106.

Piro, S., Maurello, P. and M. F. Cammarano 2000. Quantitative integration of geophysical methods for archaeological prospection. Archaeological Prospection 7 (4), 203-213.

Richards, J.D. and D. Robinson. 2001. Digital Archives from Excavation and Fieldwork: Guide in Good Practice (Second Edition). Oxford.

MOLAS. 1994. *Archaeological site manual* (Third edition). London.

Rains, M. 1995. Towards a computerized desktop: the Integrated Archaeological Database System. In *CAA1994: Computer Applications and Quantitative Methods in Archaeology*: 207-210.

Ryan, N.S. 1985a, 'Some Thoughts on an Archaeologist's Toolkit', in E. Webb (ed.) *Computer Applications in Archaeology, 1985*, pp 126–132, University of London Institute of Archaeology, London, UK.

Ryan, N.S. 1988, 'Browsing through the Stratigraphic Record', in S.P.Q.Rahtz (ed.) *Computer and Quantitative Methods in Archaeology, 1988*, pp327–332, British Archaeological Reports, Oxford, UK.

Ryan, N.S. 1995, 'The Excavation Archive as Hyperdocument?', in J.Huggett and N.S. Ryan (eds.) *Computer Applications and Quantitative Methods in Archaeology, 1994*, pp 211–220, British Archaeological Reports, Oxford, UK.

Ryan, N.S.. 2001. jnet, a successor to gnet. In W. Borner, ed., *Archaologie und Computer, Workshop 6*. Forschungsgesellschaft Wiener Stadtarchaologie, November 2001.

Silliman, S. W., Farnsworth, P. and K. G. Lighfoot 2000. Magnetometer prospecting in historical archaeology: evaluating survey options at a 19th-century Rancho site in California. Historical Archaeology 34 (2), 80-109.

Tsokas, G. N. and R. O. Hansen 2000. On the use of complex attributes and the inferred source parameter estimates in the exploration of archaeological sites. Archaeological Prospection 7 (1), 17-30

van Leusen, M. and N.S. Ryan, in press, ‘Educating the Fieldwork Assistant’, in G. Bürenhult (ed.) CAA2001: Computer Applications and Quantitative Methods in Archaeology, Visby, Gotland, April 25–29, 2001.

Wilcock, J.D. 1975, 'Archaeological Context Sorting by Computer', *Computer Applications in Archaeology, 1975*, pp 93–97.

Zdanovich, G., Kochnev, V. and B. Punegov 2003. The experience in applying 3D technologies of magnet fields interpretation got at the archaeological site “Arkaim”. Computer Applications and Quantitative Methods in Archaeology (Enter the Past, Vienna conference 2003.

8. Visualisation and rendering

8.1. Systems, tools and technologies

8.1.1. Introduction

The subject of this report is ‘visualization and rendering’ which, in a strict sense, refers only to image generation. There are two fundamentally different ways to present computer generated imagery, namely films and interactive 3D. The first implies that images are rendered offline. This is less of a problem since almost all modeling tools incorporate a (nearly) photorealistic renderer. In practice, though, there are two other issues that are more difficult to accomplish:

(a) the creation of the models that are to be displayed, and
(b) to assemble compelling virtual wolds for interactive rendering.

The latter, in turn, has also implications for modeling, and this is why both related subjects are in focus in this report. The situation for visualization and interactive rendering is characterized today by a fundamental dichotomoy, the separation of the model creation from the interactive virtual inspection. The proceeding is always basically like this:

1. Model creation – either by shape acquisition or by manual modeling for synthetic shapes
2. Export and preprocessing – to make sure the created world is amenable to interactive inspection
3. Interactive Viewing – ranging from simple 3D model inspection to rich responsive virtual worlds.

The different options for these steps and their implications will be explained in the following.

8.1.2. Model creation

A computer can generate an image only when it has something to display. This means that 3D objects need to be somehow transferred to a digital form. It makes a great difference, though, whether the goal is to create videos or content for interactive 3D. While the first can make use of all possible rendering features, and is allowed to take hours for computing a single image, the latter requires that the virtual worlds is optimized for interactivity.

In this case the computer has less than a 20th of a second (50 milliseconds) to generate one image. To achieve this usually involves decent pre-processing steps, such as global illumination, space partitioning, and a reduction of the polygon count of the model. Also one has to keep in mind that, unless the viewer motion is explicitly restricted, a model can be interactively explored from all possible sides. A model for offline rendering, though, needs to be detailed only where it is visible.

In principle two possibilities exist to obtain a digital shape:

- shape acquisition of real shapes, and
- shape modeling of synthetic shapes.

8.1.2.1. Shape Acquisition

Shape acquisition involves a whole process chain. It is only roughly sketched here.

Acquisition of raw data

This is to scan or digitize an existing real 3D model in order to bring it into the computer. This requires acquisition hardware, which can be affordable as well as expensive:

1. laser-range scanning
2. structured light approaches
3. photogrammetry

Alternative approaches include more 'exotic' techniques like computer tomography (MRT,CT), and taking explicit surface samples, e.g., with the phantom device and other measuring systems. – This subject is treated in detail in a separate report.

Processing of raw data

The result of the acquisition step is in any case a large number of points floating in 3-space. In order to convert the individual scans into a triangle mesh they need to be (a) registered and (b) post-processed. Registration means that scan points in surface regions where two or more scans overlap must be made to agree on a single surface. Post-processing is needed to remove scanning artifacts: holes in the surface from outliers, or falsely closed parts because of occlusion.

The result of scanning and shape acquisition is in any case a dense triangle mesh, in some cases even *very* dense. Further post-processing steps include simplification and the conversion into a multi-resolution representation.

8.1.2.2 Shape Modeling

To create an appealing shape with a modeling tool is generally a tedious task, and it requires specially skilled and trained personnel. In some cases it is the only alternative, i.e., when scanning is not possible. A shape might be either not be amenable to scanning because it is buried, too large, too deteriorated, or even destroyed, or it may even be just a hypothetical shape that belongs to a whole, maybe animated, series of hypothetical shapes.
One of the great advantages of using computer graphics is that authentic findings and historic artifacts can be put into a context, or even multiple contexts. The context is suggested and put to life exclusively by graphical means. It is good style, though, to clearly distinguish between fact and hypothesis, e.g., by using transparent objects or false colors. To make use of this advantage requires to create the hypothetical shapes in a 3D modeler.

Procedural modeling tools

The first category of 3D modeling tools are the fully fledged allround software packages. They are specialized in creating surface models, and targeted at creative 3D shape artists. This growing species has originally emerged from the computer games industry; appealing shapes are increasingly used in other commercial contexts as well.

The name 'procedural modeling tools' comes from the *forward modeling style* that all these tools support: The artist can choose from menus with usually *many* different modeling tools and operations that are successively applied to a shape.

The shape is iteratively changed until it matches the specifications and the ideas of the artist. This style of work requires that the artist is well acquainted with the navigation, the selection modes, and the modeling tools available from his or her chosen software. More complicated shapes also require some planning, which parts are to be created first, then refined, etc. Modeling and sculpting are crafts that simply require much experience. To some extent, experience can also outweigh the deficiencies of a modeling tool: Rather than switching over to the optimal product in every given moment, artists appear to prefer sticking to their favorite modeling software because they know it so well.

High-end CAD

CAD software differs from procedural modeling tools in that it provides very high, and *guaranteed* precision, and it integrates seamlessly with the industrial workflow in the context of computer-aided manufacturing (CAM); another distinct high-end feature is *product data management* (PDM).

Much more interesting for cultural heritage is another feature of high-end CAD tools: All of them support *parametric design*. This means that variants can be much easier produced.

Parametrics are the basis for ‘intelligent’ 3D objects, sometimes also called associative design: Every created part knows about its dimensions and its relation to other parts, which permits to re-do the same construction in many different ways. – So if it is accessible, using CAD is preferable.

8.1.3. Visualization and Rendering

8.1.3.1 Offline Rendering: Animated films and videos

The great advantage of pre-rendered films is that the author has complete control about what the audience sees, and about the moment when it is seen. The author/director can shoot his takes from perfectly chosen views, and show and explain animated processes with perfect timing. The disadvantage is that the audience might feel that movies are better seen in a cinema rather than in a museum. A movie may be boring for one and not enough detailed for the other person.

The ‘fully fledged’ procedural modeling tools all bring in facilities for defining animations, and they all contain built-in offline renderers, usually some sort of raytracer. This permits, in principle, for creating whole animation films within each of these tools. Professional users, though, prefer to use different tools in different stages, especially for the video post-processing (e.g., Photoshop, gimp...) and cutting (e.g. Premiere etc.), as well as audio dubbing and encoding.

Under no circumstances should the artistic skills be under-estimated that are required for creating videos. Public expectations are very high, and animated short films created, e.g., by computer scientists tend to be immediately recognizable as such. The same is true for other disciplines as well; more and more high schools, however, have an animation department, or at least a small video studio and a film class.

8.1.3.2. Online Rendering: Interactive 3D

The advantage of interactive presentations is that users can choose the depth and length of explanation according their personal preferences. Furthermore, interactive presentations are more engaging, and they keep the level of involvement higher; at the risk, of course, that technicalities may distract from the content. The disadvantage is that authoring of good interactive presentations is very demanding, since every possible aspect of the interaction must be anticipated by the author. Another disadvantage is that an interactive cultural heritage presentation might always be compared to 3D computer games, but the first need to be produced with only a fraction of the budget of the latter.

The requirements for interactive and real-time 3D on the technical side are also demanding. They are more restrictive than those for offline rendering. This is true for the content as well as for the runtime environment. The content needs to be specially designed for interactivity, which is possible since many of the modeling tools now support *low-poly modeling*.

Commercial 3D Presentation Environments

Pre-fabricated solutions for the professional presentation of virtual worlds indeed exist. They are mainly targeted at (a) product design and (b) VR training applications. Externally created models are imported into an authoring application, sometimes via special plugins. It distinguishes between static environment and dynamic objects, and it permits to define the motion of the dynamic objects. Things may move, with motion triggered by time (animations) or events ('behavior'). A behavior graph is basically a network of event callbacks. Motion patterns include also crowd simulations etc.

The great advantage is of course that the authoring tools are optmized for productivity. They follow an 80:20 approach so that a good-looking, fluent virtual world with standard behavior is created with a few mouse clicks. Only to walk about is considered boring by most users after a short period of time, though. Interesting VR requires an engaging plot with a more complicated story book and, thus, a complex event network. Furthermore it is not easy to realize content with a minimum of didactic and educational value without programming.

Software for commercial 3D presentations can be very expensive. Other issues include the deployment of a CH-VR experience, whether it is to be disseminated in an open fashion (CD, download) or only for a defined setup (kiosk).

Game Engines

The high-end of interactive 3D is marked by computer games. The leading handful of cutting-edge computer games are published primarily to promote the respective underlying *game engines*, which are the main source of revenues of the companies. Game producers can license an engine as the run-time infrastructure of their own games to then focus only on creating the content of their own respective games. Consequently, the costs for licensing a top game engine are typically immense, in the range of several hundred thousand Euro or USD, and definitely not affordable for CH productions.

The game companies use to license their previous engines for a much smaller fee, though, and sometimes give them even away for free (for non-commercial use). Disadvantageous are the authoring applications: They are tuned for highly skilled, professional users that usually wish to tweak and tune each individual aspect of the virtual world they create. One *can* do everything with this engines, but one sometimes also *has* to do everything. So games engines can be recommended only for very advanced content creators. Another disadvantage of games is that they are usually optimized for only very schematic types of interaction, usually to kill virtual enemies. This makes it hard to realize content with educational value.

The environment is very static, and processes and developments are not easily illustrated with a game engine. Another, very serious, drawback is the limited *sustainability* of all the immense amount of work that was needed to create the compelling content – given that the turnaround times in the games industry are between six months and two years. After two years, the look and feel of a 3D computer game is usually perceived as being obsolete.

3D Engines, Open Source and Commercial

The number of available 3D engines is absolutely amazing. The 3D engines database lists around 200, both commercial and free, and sourceforge lists around 900 projects with 3D rendering and 400 with 3D modeling in them. The different projects greatly differ in purpose as well as in degree of maturity. Engines exist that can be accessed on a high level over a scripting language (python, java), others are just software libraries to facilitate the use of the two existing low-level 3D graphics APIs, DirectX and OpenGL. It is therefore up to a decent comparison of the available features, the supported import formats, the suitability for the purpose in the context of CH presentations, etc., to answer the question of which 3D engine to choose. To describe all pros and cons is beyond the scope of this report. One drawback all engines have in common is that, like with game engines, it requires in any case substantial effort to achieve good results, and long-time sustainability is definitely also an issue. But in any case to review the current state of affairs in 3D engines is absolutely worthwhile for any interactive 3D project for cultural heritage: Given a specific purpose the perfect engine might exist.

8.1.4. Augmented reality

The ARCHEOGUIDE project provides a lot of information about augmented reality applications in cultural heritage. The deliverable on ‘prospects of augmented reality in the European Cultural Heritage sites’ is particularly interesting (http://archeoguide.intranet.gr/papers/D09_all_16.doc).

8.1.4.1. Approaches for Augmented Reality

To be able to augment computer generated information into the user’s view of the real world, thus to match the computers virtual world into the real world, a set of different parameters is necessary. These parameters describe on the one hand the user’s position and orientation with respect to the scenery and on the other hand the position and orientation of the scene within a global world co-ordinate system.

In order to match computer generated information into the user’s view, the user’s position and orientation is needed. To get the information, tracker systems are used, that report the locations of the users and the surrounding objects in the environment to the augmented reality system. Most commercial tracker systems were developed with respect to the needs of virtual reality environments, and are not suited for augmented reality environments. Depending on the application area, different tracker systems and sensors are in use like Global Positioning Systems (GPS), that are used for outdoor tracking, ultrasonic tracker, usually used in indoor application scenarios, or inertial tracker. Recently, research has been performed focussing on tracking solution using hybrid tracking technology.

Besides the tracking, which delivers some basic information on how to augment computer generated information into the user’s current view, there are different approaches on how to display the information. The most typical approach is to use head mounted displays, that are either *optical* see-through HMDs or *video* see-through HMDs. Depending on the application scenarios, other approaches use e.g. hand-held video see-through displays.

8.1.4.2. Challenges of Augmented Reality

One of the biggest challenges in the area of Augmented Reality is the *registration issue*. To create a sufficient illusion of the augmentation, it is necessary to properly align real and virtual objects with respect to each other. Otherwise the illusion that the two coexist will be compromised.

For a number of reasons, the problem of registration is difficult to solve. First, the humans visual system is very sensitive for even small misregistrations. Second, the errors, that were indeed tolerable in Virtual Environments are no longer acceptable in Augmented Reality, so misalignments in HMD, error of the head tracking systems as well as other problems have obvious visible effects in Augmented Reality scenarios. Third, there is still a problem of system response time: the time interval between tracking and the superimposing of the computer images on the real world. The effect is that the virtual objects may "lag behind" their real counterpart as the user moves around.

The problem of registration was addressed in various research projects, but because of the problems nature, solutions successfully applicable to one application scenario are not naturally usable in a different application scenario.

Nowadays research project in the area of Augmented Reality do not address the issue of augmentation, but keep track to more application oriented scenarios, as the user's ability to interact with the real and virtual environment has become more important.

8.1.4.3. Application areas and current projects

Currently, a large number of research institutes and groups is facing the challenges of Augmented Reality when considering different applications areas. To name a few, in Germany the Fraunhofer Institut for Computer Graphics (Fraunhofer IGD), Center for Computer Graphics (ZGDV), the University of Munich, the University of Hannover , and many more are doing research the field of Augmented Reality. In the US, the MIT, the HRL Lab, Carnegie Melon University, Columbia University, Rockwell Research and for Japan Sony Computer Scientific Laboratory, Mixed Reality Systems Laboratory are examples. Interestingly, even though Augmented Reality is a very promising technology with the potential to be applied to any different application areas, very limited transfer into real applications is noticeable.

The application areas, current projects are focussing on are widespread, ranging from medical applications, assembly and maintenance applications in various engineering fields to applications scenarios in the field of entertainment and education as well as cultural heritage like ARCHEOGUIDE. The specific research focus of the different projects regarding tracking and registration tightly depends on the requirements of the project's intended application scenarios, e.g. some deal more with tracking or registration issues in closed or open environments, some deal with the issue interaction in Augmented Environments.

8.1.4.4. Some applications of augmented reality in cultural heritage: virtual tours

The Ename 974 Project[72]

The Ename project uses a computer technology called TimeFrame in order to provide on-site 3D reconstructions of monuments in archeological sites and allow the presentation of sites to different audiences without the physical reconstruction of their remains. The TimeFrame consists of a fixed video camera, a computer system, two monitors and a touch screen. A booth or “kiosk” protects the system and the visitors. The camera is directed at a particular section of the archeological site and transmits real-time pictures to the monitor screens. The system superimposes the real scene with 3d reconstructions of the monuments and displays the resulting scene to the monitor allowing the visitors to visualize how the monuments appeared when standing.

The ARCHEOGUIDE project[73]

ARCHEOGUIDE (Augmented Reality based Cultural Heritage On-site GUIDE) is an EU-IST funded project that will provide the users the experience of a tour in a cultural site with the ability to view the natural environment, to visualize 3D reconstructions of monuments and be assisted during the view by a multimedia guidance system. The ARCHEOGUIDE system will address the requirements of a wide user selection that includes cultural site visitors, cultural site managers, researchers, and content creators. Cultural site visitors will be provided with a see-through Head-Mounted Display (HMD), earphone, and mobile computing equipment. A tracking system will determine the location of the visitor within the site. Based on the visitor’s profile and his position, audio and visual information will be presented to guide and allow him/her to gain more insight into relevant aspects of the site.

[72] http://www.ename974.org

[73] http://ARCHEOGUIDE.intranet.gr/

8.2. Future directions and EPOCH's contribution

The most significant future direction of interactive 3D with respect to cultural heritage is the development of standards. Institutions dealing with CH usually also deal with history; and this implies a certain attitude towards issues such as sustainability and long-time archival, but also towards cost efficiency.

The development of standards will be the way out of a serious dilemma: 3D technology is currently too expensive for museums and CH institutions. Resources are there to use it only for demonstrations and pilot projects, but not on a regular basis. On the other hand, museums will more and more depend on *interesting 3D*. Its potential advantages are manifold:

Historic relations, developments, processes can be made very, very clear with animated three-dimensional illustrations. Real findings and artifacts become more interesting when their use and historic context is shown in an appealing and detailed way. Museums are conservative by nature, but to gain attention they have to compete with theme parks and fun experiences, if they are willing to accept this or not.

3D must mature further.

Affordable high-end 3D hardware is available only since around the year 2000. Ever since it was the games sector that has pushed the development; the significance and usefulness of 3D goes far beyond games, though. The situation today is that there is not 3D hardware-, but a *3D software crisis*.

A prerequisite for a fundamental advancement of 3D technology is that 3D applications must become *interoperable*. Note that an interactive 3D presentation created by the 3D presentation environment Quest3D is usually deployed as an executable binary. This alludes to the fundamental problem that an important constituent of an interactive 3D experience is determined by procedural aspects, by the look-and-feel, by the temporal and event structure, the animations etc. A 3D experience is *executed* rather than consumed in a strictly sequential fashion like a text or a video. – Only the claimed interoperability will permit to increase productivity and cost-efficiency, e.g., by exchanging 3D templates and even whole interactive exhibitions between museums and CH institutions. This, in turn, will foster a new market for 3D experiences, a cultural heritage edutainment industry that does not yet exist today.

Second, 3D presentations must become *more interesting*. Most currently available 3D presentations do not at all exploit the full potential of 3D. As renowned 3D professionals use to put it: "You ain't seen nothing yet" (John Carmack, id software). It appears that completely new artistic and design approaches may need to evolve; one indication are new form of art such as, e.g., machinima.

There is also a very concrete technical reason for the fact that 3D today is used insufficiently, and also in unsatisfactory ways. One inherent limitation of 3D today is the *separation* *between modeler and viewer*. Highly sophisticated modeling tools are used to create the most decent virtual objects, but only through a (long) sequence of manually executed processing steps. Then all intelligence is removed from the 3D objects, as they are exported to a dumb low-level triangle mesh: The shapes lose all their semantic information, their modeling history, their changeability, all that remains is a ‘frozen’ mesh.

From this are even a lot of triangles removed then (low-poly version), and the only thing left the poor resulting objects are good for is to move, rotate, and scale them, and to instantiate them several times.

With respect to this fundamental limitation the aforementioned variety of 3D engines and presentation formats is an illusion. What is common to *all* 3D presentation environments is that they only permit to *move* things around, and not to *change* things.

8.3 References

8.3.1 Procedural modeling tools

There is a limited number of mainstream modeling tools. To some extent the market is dominated by the two high-end packages *3D Studio Max* and *Maya*. Some of the modeling tools with wider spead use are listed in the following list. They can be compared through downloading trial versions. It is worthwhile to have a decent look at their feature lists: Most artists admit that in practice they use only a small fraction of their favorite tool’s modeling functionality.

Maya 2099 Euro www.alias.com
3D Studio Max 3500 Euro www.discreet.com
Houdini 1300 USD www.sidefx.com
Rhinoceros 900 USD www.rhino3d.com
Cinema4D 700 Euro www.maxon-computer.com
Lightwave 600 USD www.newtek.com
Caligari truespace 600 USD www.caligari.com
SoftImage XSI 450 Euro www.softimage.com
Sketchup 475 USD www.sketchup.com

Remark: The last tool, Sketchup, is the author’s personal recommendation for a 3D sketching tool for casual users. Sketchup is ’lean and mean’, it is easy to learn and use, and it yields very clean polygonal shapes.

8.3.2 High-end CAD

The high-end CAD market is firmly dominated by the ’major CAD companies’. They provide high-end functionality at a high-end price. But they usually have very good conditions for academic and educational institutions.

AutoCAD www.autodesk.com
Dassault Catia www.3ds.com
Pro/Engineer www.ptc.com
SolidWorks www.solidworks.com
UGS SolidEdge www.solid-edge.com

8.3.3 Open Source Modelers

OpenCascade www.opencascade.org
Blender www.blender3d.org

8.3.4 Commercial Game Engines

Halflife www.valvesoftware.com
Quake www.idsoftware.com
Unreal www.epicgames.com

8.3.5 Overview Pages

sourceforge.net – Sourceforge Open Source Software Repository

The number of graphics related software packages on sourceforge alone is 924, of which 901 list '3D Rendering' among their keywords: sourceforge.net graphics software.

www.devmaster.net/engines – DevMaster's Game and Graphics Engines Database

Database devoted to 3D engines, commercial and open source, currently featuring 197 detailed entries, among which are:

Irrlicht very active, zlib/libpng license irrlicht.sourceforge.net
Cube freeware first person shooter www.cubeengine.com
Codecult commercial realtime engine www.codecult.com
RealityEngine commercial realtime engine www.artificialstudios.com

www.3dlinks.com – Ultimate 3D Links

Moderated overview about 3D software and plugins. User community sorted by tools. ***Very useful***.

www.cgtalk.com – Creative Computer Graphics Forum

Meeting point for practitioners, interesting tricks and howtos, good starting point for questions and research.

9. Multi-modal interfaces

9.1. Systems, tools and technologies

The hardware interfaces used for Virtual Reality applications are very different from those used in traditional desktop interactive systems. Hence, there is not only one good solution but instead many different hardware technologies which suitability, performance and cost depend on the specific type of application. This document attempts to identify the available hardware technologies for assembling an efficient and cost effective virtual reality installation. In particular, these technologies are suitable for presenting real time and interactive cultural heritage virtual environments. These technologies are classified as follows:

1. Output devices
   - Visual displays technologies
   - Audio systems
   - Haptic and tactile systems

2. Input devices
   - Interaction devices
   - Motion capture

9.1.1. Output devices

These technologies present information to the user/users of the virtual reality installation. Typically, three sensorial modes of input information are used: visual, sound and tactile information.

A. Visual display technologies

Currently, the display industry is a fast growing dynamic market with new technologies being constantly developed. It is therefore difficult to find a single leading technology suitable for all types of applications. In order to identify the major current technologies available for efficient and cost effective virtual reality installations, the following subsections will present the main commercial display technologies and their characteristics. These characteristics (resolution, contrast ratio, brightness, display size and shape as well as stereo capabilities) will be used to compare the different technologies and they are explained in more detail in the following subsection.

A.1. Display technology characteristics

Resolution

Resolution is one of the most important factors to consider in any display technology. It refers to the amount of detail that can be seen on the screen and it is usually described by the number of columns and rows of pixels that can be displayed (NEC/Mitsubishi 2005). Many displays and projectors have fixed resolutions and ratios.

The key to a quality image is the best resolution that can be afforded (Musgrave 2001). According to Parkhurst and Niebiur (2002) a resolution of 18 mega pixels should be feasibly for large scale displays.

Another important issue to consider is that as the resolution increases, the performance of the graphic card will drop. According to Pabst (2000) and Thon et al. (2004), the memory bandwidth of the GPU (Graphic Processor Unit) has a big impact on the frame rate at 16-bit colour depth and high resolutions (see figure 6). This gets worse at 32-bit colour depth because the amount of data that needs to be transferred between the 3D-chip and the local memory doubles almost exactly.

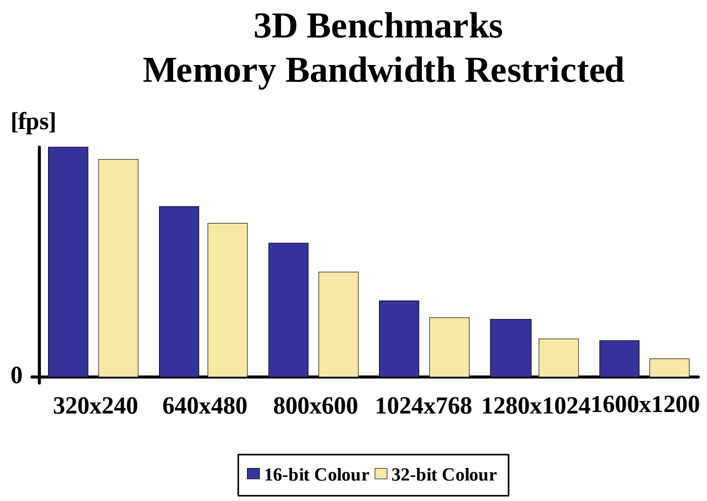

Figure 6. GPU memory bandwidth restrictions (Pabst 2000 and Thon et al. 2004)

Contrast Ratio

Contrast ratio is the measure of the difference of brightness levels between the brightest white and the darkest black. This is a very important specification in judging how well an image will be produced in an environment with different ambient lighting (Musgrave 2001). A good contrast ratio of 120:1 easily displays vivid colours, while ratios as high as 300:1 support superior grey scaling (NEC/ Mitsubishi 2005).

Brightness

Display brightness, which is measured in nits (candelas per meter squared cd/m2), can make a significant difference in a user's visual experience. A higher number of nits indicate a brighter display. As a general rule, 1,500 to 2,500 nits for an indoor display are required and 5,000 or more for outdoor displays (Musgrave 2001).

Display size and shape

Visual displays with different sizes and shapes can provide different degrees of immersion in a virtual environment. The degrees of immersion are usually measured by the field of view (fov), which describes how much of the user's view can be covered. Kjeldskov, J. (2001) suggests a classification of displays technologies according to their degree of immersion. These are:

1.) Full immersive displays: supports the feeling of "being in" the virtual environment
    a. Head Mounted Displays (HMD)
    b. Booms (HMD mounted on stand and operated by hand)
    c. Six-sided Cave
2.) Partial immersive displays: supports the feeling of "looking at" the virtual environment
    a. Large monitor
    b. Panoramic screens
    c. 3-5 sided caves

Stereo capabilities

Stereo means the capability of viewing images on the display in 3 dimensions as they are naturally perceived by the eyes. Some of the displays technologies offer stereo capabilities, which can be achieved as follows:

- Active Stereo: the viewer wears special eyewear consisting of two controlled shutters working in synchronization with the projector (see figure 7). When the projector displays the left eye image, the right eye shutter of the active stereo eyewear is closed, and vice versa. The single projector used must be capable of displaying at a refresh rate to alternate high enough that the viewer does not perceive a flicker between alternate frames (Kjeldskov 2001).

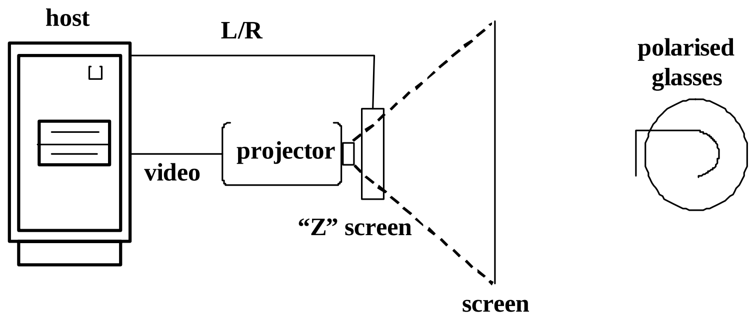

Figure 7. Active stereo display (Belleman et al. 2001)

- Passive Stereo: the viewer wears a pair of glasses containing two oppositely polarized filters -one for the left and one for the right eye. The light from each of the two projected images is polarized and can only pass through its corresponding filter. It can be implemented using one or two projectors, as shown in figure 8. The two-projector approach has the added value of providing higher brightness (Kjeldskov 2001).

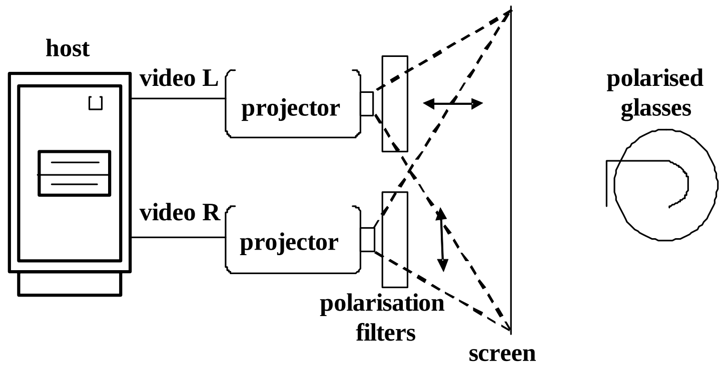

Figure 8. Passive stereo display with two projectors (Belleman et al. 2001)

- Auto stereo: this type of display allows the 3-D image to be viewed without glasses from a wide range of angles and distances. They are a relatively new addition to the 3D display market and open up a wide range of new opportunities to exploit 3D imagery. However, new problems are also introduced with the technology. For example, rather than displaying two distinct views, the display provides up to 9 which introduces new demands on other characteristics, such as resolution (Inition 2005).

Cost

Cost is a major factor to consider when deciding which display technology to use. As there are big differences of cost between display technologies, the available budget could become an important determinant for selecting a specific technology. Moreover, it is important to notice that prices of visual display vary rapidly due to the very dynamic nature of this industry.

A.2. Display technologies

Some years ago CRT (Cathode Ray Tubes) displays were the single prevailing display technology available in the market. Today CRT, LCD (Liquid Crystal Display), PDP (Plasma Display Panel), DLP (Digital Light Processing), and LCoS (Liquid Crystal on Silicon) are mature and mainstream, while many other technologies, such as OLED (Organic Light Emitting Diode) and FED (Field Emission Display) are trying to emerge from the development lab and grab a significant market share.

An attempt to classify these technologies can be made by dividing them in direct view or projected technologies. Projected technologies cover both front and rear projection as the technology is shared between both markets. Figure 9 shows this classification, where CRT and LCD technology can work with either method. On the other hand plasma is limited to direct-view while DLP, LCoS and D-ILA are limited to projection because the devices are small microchips.

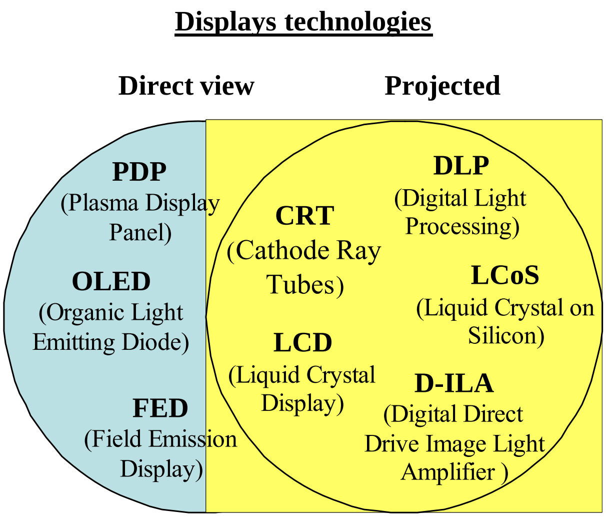

Figure 9. Display technologies classification

These technologies will be discussed in the following subsections:

Cathode Ray Tubes (CRT)
This 75 year old technology benefits from excellent colour and supports a wide range of resolutions. If overall image and picture quality is the absolute criterion then the direct-view CRT displays remains the best option, especially for video. However, sharpness and contrast aren't as good as the flat panels operating at their native resolution and they are bulky and heavy (Soneira 2005). Therefore, according to DeBoer (2004) this technology is already extinct as the new display technologies offer better capabilities and fewer people are willing to pay a premium price for a CRT.

Plasma Display Panels (PDP)
Plasma displays have captured a significant share of the non-CRT video market, mainly because their screen sizes (up to 80 inches) and lows prices. As plasma technology does not use backlight and no projection of any kind, they achieve sharp and high contrast displays with excellent colour and wide viewing angles (DeBoer 2004). However, they are susceptible to the “screen door” effect due to noticeable gaps between pixels, have fixed native resolution and are very heavy. The screen door effect is when the space between pixels is visible as a minute black box around the pixel. The performance of this technology has been steadily improving with size and brightness. The real question is how plasma displays will hold up to the challenge from direct-view LCD panels.

Liquid Crystal Displays (LCD)
Currently, LCDs are the dominant flat panel technology for computer applications. They offer good colour reproduction, perfect sharpness at native resolution, they are very thin and not very heavy. Some of the disadvantages of this technology include fixed native resolutions; poor contrast ratios, brightness and colour saturation generally decrease as the viewing angle increases and difficulty to produce deep blacks. However, this technology is increasingly

getting bigger, faster and cheaper and many analysts predict that they will eventually take over the plasma market (DeBoer 2004, Soneira 2005).

Active Matrix Liquid Crystal Display (AMLCD)
High Temperature Poly-Silicon (HTPS) is an example of this variant of LCD technology. Its advantages over LCDs are that it is smaller, has higher resolution and higher contrast. Its main disadvantage is that pixels cannot be made smaller without lowering brightness and reducing picture quality. This means that to produce enlarged panels without shrinking each pixel results in extra cost (DeBoer 2004).

Reflective LCD
This technology is a variation of LCD, where a liquid crystal layer sits on top of a highly reflective substrate, instead of being embedded in the LCD material as with conventional LCD technology. This assembly used in a projection subsystem, providing an excellent contrast ratio (Musgrave 2001). Liquid Crystal on Silicon (LCOS) and Digital Direct Drive Image Light Amplifier (D-ILA) are examples of this technology and they will be described in the following subsection.

(a) Liquid Crystal on Silicon (LCOS)
LCOS is a reflective LCD display panel, which can be created as 1 chip or 3 chip systems. Pixels on LCOS panels can be made smaller than is possible with other microdisplay technologies, without compromising picture quality or manufacturability. Other advantages of this technology are excellent colour reproduction, high resolutions and no "screen door" effect. The main disadvantages are their cost and the difficulty to access this technology (DeBoer 2004). LCOS technology is still relatively expensive compared to LCD and DLP projection technology. However, it is predicted that it will be developed to become better, faster and cheaper with Intel stepping in.

(b) Digital Direct Drive Image Light Amplifier (D-ILA)
The D-ILA device, a special LCOS technology developed by JVC, is a reflective type of LCD that delivers a greater amount of light than a transmissive LCD panel. This technology eliminates almost entirely the "screen door" effect leaving a nearly-seamless picture. Overall, D-ILA projectors produce higher resolutions, better contrast ratios, less image artefacts, and better tonal and colour information than LCD front projection devices. However, this technology is very expensive and it is not better than DLP projectors (DeBoer 2004).

Organic Light Emitting Diode (OLED) Displays
OLED technology is based on a revolutionary discovery that light-emitting, fast switching diodes could be made from polymers as well as from semiconductors. OLED technology provides not only brighter, better images at a lower cost, but it uses a material with self-luminous properties that eliminates the need for a backlight (DeBoer 2004). This means that they draw far less power and they can be used with small portable devices. Other advantages of this technology are that the manufacturing process is simple and that there is virtually no restriction on size. There are two main directions in OLED (Howard 2004):

- The first technology was developed by Eastman-Kodak and is usually referred to as "small-molecule" OLED.

- A second technology, developed by Cambridge Display Technologies (2005), is called LEP or Light-Emitting Polymer.

This technology is quickly emerging from the development lab. Some of the technical problems left to overcome before this technology definitely enters the display market are its lifetime, ensuring competitive refresh rates, contrast ratios, black levels and overall performance (DeBoer 2004).

Field Emission Display (FED)
The biggest advantage of this emerging technology is that the display is emissive, and thus has an excellent viewing angle (160 degrees both vertically and horizontally), while remaining as thin as an LCD. In addition, they are also several microseconds quicker in response speed (Sharp 2005). Since the FED is a vacuum device, atmospheric pressure becomes a severe problem for large-area panels. This together with lifetime issues and bringing down the driving voltage are the main challenges ahead for FED developers.

Digital Light Processing (DLP)
Since its commercial introduction, DLP has steadily improved performance and now it is a dominant technology for both computers and video. This technology uses a semi-conductor imaging chip developed by Texas Instruments called a DMD (Digital Micromirror Device), which can be made up of one or three chips. DLP projectors offer excellent colour reproduction and produce a smooth yet sharp image with no apparent pixelation except close to the screen (Musgrave 2001). However, they have fixed native resolution and they can cause possible visual fatigue due to temporal fickle and rainbow effects in single chip systems (Soneira 2005). This technology is only available for rear projection and authors agree its future challenge will be to compete with the direct-view display market (DeBoer 2004).

A.3. Direct view vs. projection displays
Direct view and projection display technologies are intensely competing in the market. Table 1 summarize the characteristics for both direct view and projection technologies according to DeBoer (2004), Soneira (2005), CTL (2005) and Howard (2004). These characteristics overview is further examined in order to compare all the presented technologies according to their main characteristics.

Displays	Max. Resolution	Contrast Ratio	Brightness	Screen size
Direct View				
CRT	1920 x 1080	4000+:1	1000 cd/m2	20" - 40"
PDP	1366 x 768	3000:1	700 cd/m2	30" – 80"
LCD	1280 x 1024	1300:1	450 cd/m2	1" – 57"
OLED	143 pixel per inch	1000:1	200 cd/m2	- 20"
FED	1280 x 768	1000:1	250 cd/m2	2" - 40"
Projected				
CRT	1920 x 1080	5000:1	NA	42" - 65"
LCD	1280 x 1024	800:1	450 cd/m2	42" – 70"
DLP	1280 x 720	5000:1	750+ cd/m2	43" – 61"
LCOS	1920 x 1080	2000:1	750+ cd/m2	42" – 80"
D-ILA	2048 x 1536	1500:1	550+ cd/m2	NA

Resolution

There is not a very large difference between the resolutions achieved by the direct view and the projected displays. D-ILA, LCOS and CRT technologies offers at the moment the higher resolution. In terms of image quality, direct view technologies are generally sharper because the pixels are generated right at the screen, while the optics and screen reduce sharpness in projection units. However, direct view technologies show much greater pixilation because the red, green and blue elements are tiled side-by-side on the screen, which means that the “screen door” effect is much greater than in projection technologies. This effect doesn't apply to direct-view CRTs because the phosphor elements are generally much smaller than the pixel size. This effect decreases as the resolution increases because the eye is unable to resolve the pixel structure at normal viewing distances.

Contrast

Regarding contrast, both direct view and projection technologies offer different contrast ratios. The best contrast ratios are achieved by DLP and CRT technologies, followed by plasmas displays. LCD projection offers the lowest contrast ratio.

Brightness

Direct view displays are much brighter than projectors. According to Musgrave (2001), the main limitation of projectors is precisely their brightness. A typical large-venue video projector with 750 cd/m2 of output will achieve only about 37% of the required brightness level.

Display size and shape

The size and shape of these displays, regardless of being fully or partially immersive, differ largely. For example, head mounted displays (HMD) and booms are fully immersive and just require a small display. On the other hand, panoramic screens and caves of 3 to 5 sides require larger displays and are only partially immersive. This suggest that the degree of immersion is not fully related to the size and shape of the display but rather to how much of the user’s view is covered by the display. For full immersion, smaller displays are positioned closer to the user's eyes. Nevertheless, this solution provides the disadvantage that only one user can be immersed in the environment at any time. In order to immerse several users, larger screens are required although this solution only offers partial immersion. For this reason, there is not any predominant display solution in the market, as there is always a compromise between degree of immersion and number of users. Therefore, the suitable display's shape and size depends greatly on the type of application and the preferred number of users to use simultaneously the virtual reality installation.

In general, direct view screens are suitable for displays of smaller sizes, such as head mounted display, boom and large monitors. On the other hand, projection displays offer better image quality for larger displays sizes. Tiling screens or projectors is usually used to achieve larger display areas. This is because many display technologies cannot produce an area larger than 80". As a result, it is necessary to address additional hardware and software problems, such as geometric misalignment, colour variation (Majumder and Stevens 2005) and other extra requirements imposed on the video controllers, which drive the display.

In terms of thickness, LCD and plasma direct view displays' main advantage is that they are very thin, typically 3 to 5 inches deep. In contrast, rear projection requires more space in the room, while front projection could be overshadowed by the users in front of the screen.

Stereo capabilities

As previously explained, stereo projection could be achieved by using projection technology. As such, CRT, DLP, LCD, D-ILA and LCOS projectors are commonly used for Virtual Reality installations with stereo capabilities. In addition, LCD and PDP direct view technology also offers stereo and auto stereo capabilities (Qinetiq 2005, Opticality Corporation 2005).

Cost

The cost of the different technologies varies widely depending on whether the technology is fully established in the market, its size, quality and performance. As previously discussed, direct view technology usually is used for smaller screens. Therefore, the prices of these technologies, which are very positioned in the market, are more affordable. On the contrary, very large direct view screens are in the more expensive side of the range price, especially when having stereoscopic capabilities. Projectors provide a cheaper option to develop a stereo capable large area display. However, depending on the technology their prices are very different, and in case that more than one projector is needed, this price multiplies.

A.4. Display technology conclusions

The previous subsections have described advantages and disadvantages of the main display technologies available in the current market. Regarding the direct view technologies, many experts believe the present might very well be plasma's high point and that the near future will belong to LCD (Bohannon 2004). In addition, OLED technology has to go from novelty to practical competitor in a market that is constantly evolving. Unlike the screen displays, the competition between the projection technologies is closer. Currently, DLP, LCOS and LCD are the main players competing to achieve best picture quality at the lowest cost. On the other hand, D-ILA technology offers some advantages over the other technologies, such as higher resolutions and better contrast ratios, and will be battling for the next several years, something that will undoubtedly be good for the market.

In conclusion, new technologies are constantly being proposed which indicates that at the moment no single technology can fulfil all applications (IST 2005). The best advice for selecting which technology is the most suitable for assembling a Virtual Reality installation will be to produce a detailed specification outlining and prioritising system requirements. These requirements could be whether the installation will be outdoor or indoors, fully or partially immersive, number of people to accommodate, budget and others. Subsequently, these requirements should be closely matched with the characteristics provided by each technology and display type. Other useful recommendations could be to evaluate actual displays in similar usage situations, and to check references of the supplier (Musgrave 2001).

B. Audio Systems

Before discussing the different types of setup, it is useful to discuss the types of information that can be communicated through audio.

B.1 Types of audio data

Audio data can be discrete, as in individual short pieces of sound of music, or continuous pieces of music or audio in general.

These different categories have different properties, and have different requirements for implementation within a virtual environment.

Discrete audio

Example: Sound effects, speech, simulated audio systems

Events in the real world often generate sound. As we experience life, we learn to recognise these sounds helping us to understand that an event has taken place, such as a door closing, or a gun firing. Often we chain these events together to understand them, our minds gain context from such information and can use predictive abilities to help us continue our understanding. An example of this is speech, where individual words are spoken and joined to form sentences for us to interpret.

A unique characteristic of discrete sound is that it can have both duration and a relative direction to the user. This ties directly with environmental audio systems, which can simulate directional sound.

Continuous audio

Example, background music, waterfalls running etc

Continuous sound often comes from the joining of many different discrete audio cues that recognised as a whole. Often, continuous sound can sound like it comes from an area, as opposed to a discrete position. It is the natural grouping of many discrete audio cues, but as the cues renew themselves throughout their duration, are considered continuous.

B.2 Types of audio system setup

Audio systems can be categorised by their use, either as part of an environmental setup, used by many users, or as a personalised setup, involving a single user. It is also possible to embed personalised setups within an environmental setup, and integration of these two types of system has issues of its own.

Environmental setup

An environmental sound system refers to the use of a single system for the interaction with many users. This refers to any system which is external to a group of users, and which responds to the group as a whole.

Examples of this setup could be, background music in a shop, sound effects in a theme park ride, or audio in a cinema system.

(a) Directionality of audio cues

An environmental setup is useful for communicating both discrete, directional cues, as well as continuous audio.

Directional cues are relative, meaning that if an event happens, in order to calculate which direction the user should hear it from, the both the users' position and the event position must be known. It is assumed that the user will be standing in the centre of the environmental sound setup, however with a group of people, that isn't necessarily the case, and this must be taken into consideration when determining how effective a system is in communicating a single event to multiple users.

(b) Group policy

By grouping everyone together in an environmental setup, the same information is given to everyone, and this must be taken into effect when constructing the experience – as not all information may be suitable or best used with all users of all types.

(c) Privacy

By grouping everyone together, users are not able to take control and listen to individual content without anyone else hearing it also.

Personalised setup

A personalised sound system refers to the use of a single system for the interaction with a single use at a time. This refers to any system which is external to the individual, but may have many systems within a group.

Examples of this setup could be a personal hi-fi or connection via a PC.

(a) Individual policy

As a per-user approach, each setup can be tailored for the user, such as adjusting the content, volume, style etc according to the users' needs.

(b) Exclusion

Exclusion refers to the ability to exclude external audio from the personalised sound stage, whether that is from an environmental audio system, other users, or indeed other uses' personalised audio setup. If it is desired, a personalised setup could be adapted to augment the outside world, as opposed to supposing it, meaning that environmental cues could be given to users as a group, but personalised cues could be given on a per-user level.

(c) Directional issues

The implementation of a personalised setup is currently focussed on using headphones. Headphones, by their nature, are stereo only, and therefore cannot handle direction sound except as a projected soundstage from left to right in front of the user. There is potential for simulating directional cues, either through a simulated soundstage (as demonstrated in Dolby Headphone) or through a combination of head-tracking and headphones, to adjust output.

B.3 Implementation options

With a good understanding of the theory, it is possible to assess the state of current technology to mix and match parts to fit our needs.
Our audio will be actuated using speakers; each speaker fed via a channel a signal that is to originate from a computer. Sound cards are a standard method of producing such a signal from software, no matter what type of setup.

In this section different speaker setups will be discussed to provide the two main types audio system.

Environmental setup

Environmental setups require a series of standalone speakers that together form the overall soundstage. The number and configuration of these speakers can be further categorised.

(a) Mono sound systems

The first types of system were concerned with listening to music or radio used a single speaker as sound actuator. From a technical point of view, a single speaker can only give information from a single direction – which is fine for speech, on a radio broadcast, but not entirely suitable for listening to music. It is as though the entire sound stage has been compressed through a single pipeline.

(b) Stereo sound systems

Stereo, with two speakers, can act together to produce a front-side horizontal sound stage. This is similar in sound to going to a concert or performance watched from the front, giving differentiation between different music cue positioning. However, it is only front facing.

(c) Quadraphonic and Surround sound systems

Where stereo failed to give a wider “virtual soundstage”, quadraphonic sound was designed to provide an early version of surround sound, by placing four speakers at four points around the user, usually in the four corners of a room.

With the advent of Dolby Surround sound, the virtual soundstage has expanded for use with a television screen. Dolby Surround comes in the form of four surround sound speakers for “effect”, whilst the focus of the user is upon the screen in front. A speaker is positioned directly in front of the user specifically for the user to listen to high definition speech, and a further speaker is placed somewhere in the vicinity giving actuation to low level sound.

This does not correspond perfectly with the virtual environment, which by its nature may have a focus that moves dramatically, the user possibly focussing on more than one screen, being able to look around in full 3d throughout the environment. Speech or any other effect may come from any direction, whereas the Dolby Surround system is based on a planar Virtual Soundstage, having particular issues for example with non-planar sound effects such as plane flights overhead.

It has been shown that 5.1 surround sound, and its extension counterparts 6.1 and 7.1, gives good directional sound as far as a front facing focus is concerned. It has, however, been used in some of the latest games to good effect, the most notable being the horror/action game “Doom3”, total immersion lends strongly to atmosphere, especially concerning surround sound – being able to hear things behind you before you look around! It is certainly possible, although not optimal to be able to use a surround sound system to work with the virtual environment to give good directional cues, lend atmosphere and depth to an otherwise mostly visual experience.

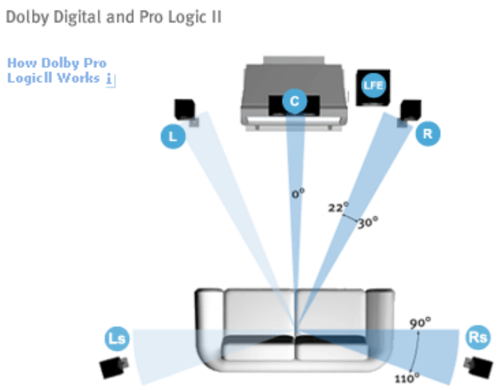

Figure 10. 5.1 Surround Sound setup using the Dolby Digital Surround sound setup

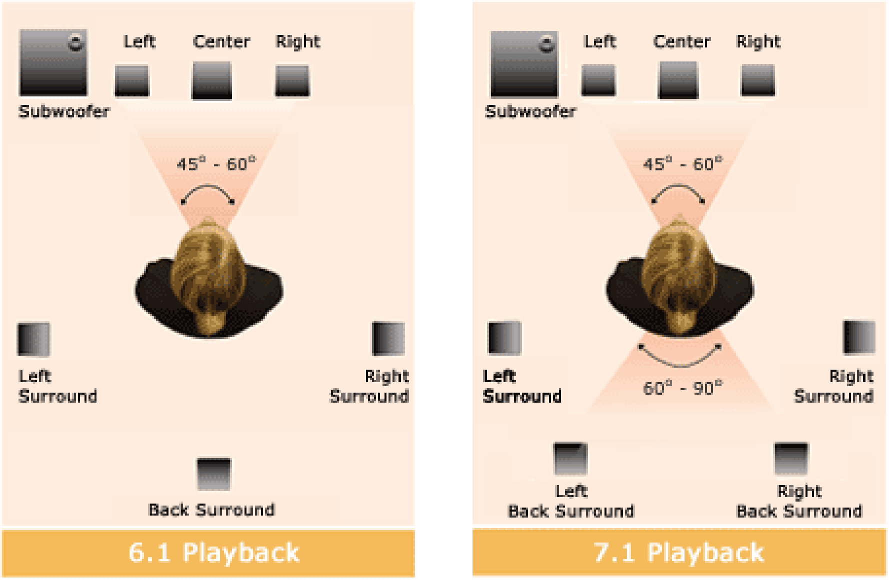

Figure 11. 6.1 and 7.1 Surround Sound setup

The speakers themselves have a wide range of prices and a wide range of quality and/or power output. Whilst the quality itself does not have a unit of measurement, power output is measured in Watts, a mini system being between around up to 10Watts (suitable for a child's' room), a midi-system being up to around 40 Watts, a living room / master bedroom size system around 100 – 150 Watts and upwards. There really isn't an upper limit for power output, although it is a concern when purchasing speakers that whether the system is specifying the Watts per speaker or of the system overall. Surround sound speakers can be in the upwards of 50 to 150 Watts per speaker, although as specified there really isn't an upper limit. Often, however, a higher power output also means that at lower output, there is better quality, and less distortion of the signal, but this is by no means a fixed property.

Personal setup

Stereo headphones come in 3 main types: internal, closed and open. Each of these types has its advantages and disadvantages, but provides sound for a single user with stereo sound.

(a) Internal Earphones

These earphones have the greatest range in price and quality, and are placed just inside the ear. They are designed to block outside sound, and generally are of lower quality than the other two, external earphone types. They are usually used with personal stereos.

(b) Closed Earphones

These earphones are vary less in quality and tend to be better than internal earphones, although sound is still slightly distorted due to their external sound blocking ability.

This ability, however, does lend to the audio environment when in a busy or noisy area, but obviously would not be suitable in an augmented audio environment.

(c) Open earphones
These earphones are usually much better quality and slightly more expensive, but do not attempt to block outside sound. This makes them better in quieter environments, and could have direct application in a virtual environment employing augmented audio, assuming that multiple people using them would not interfere with each other.

PC Output
Sound cards are a good method of facilitating the conversion of a digital signal generated through software to the speakers, either in digital format via a proxy or directly from the sound card itself. There are a range of different options concerning sound cards, varying in quality of output and many options that may not be necessary in a virtual environment, such as audio input or internal software tables for the use of midi devices.

Integration issues
For connecting speakers to a soundcard, the best quality is achieved through the use of a digital A/V receiver, which acts as a control and amplification unit. An A/V receiver is designed to accept audio from a digital source an redistribute to a multitude of different speakers dependent upon the setup.

Below is a summary of possible audio devices, implementation aspects and possible usability issues:

5.1, 6.1 or 7.1 Surround Sound Systems, based around an A/V Receiver
- Requires software implementation for 3D sound
- Multiple users can enjoy simultaneous use.
- Multiple users can enjoy directional sound.

Headphones / "augmented audio"
- Stereo only, no 3d sound
- Requires dedicated sound card, but can be cheap (>10 pounds)
- Requires software producing sound
- Individual users can enjoy customized use.

Sound cards
- Some can produce surround or stereo signal directly for connection to speaker setup.
- Some can produce optical digital output using SPDIF connector to an A/V receiver for digital quality surround sound.
- Also have variations in the quality and range of midi and recording support, which may or may not have an effect in virtual environments.

C. Haptic feedback systems
Haptic devices are concerned with the communication of physical information to a human user. There are many types of physical cue to consider.

C.1 Environmental cues

Environmental cues are communicated to a group of people as a whole and represent many types of physical phenomena that are taken for granted in the real world. Examples of these types of cue include physical movement and orientation, force feedback due to environmental effects such as wind or earthquake etc.

There are many issues concerning an environmental.

Blanket user policy

Each user within the environment will be treated in the same way, however, different users will experience the same effect differently. There are immediately issues concerning the types of user that can use a particular environment, such as the suitability due to age, disability, or other factors.

Latency and refresh rate

A simulator is an example of a virtual environment. It is a combination of a seated theatre, in which a movie is shown, with a hydraulic or pneumatic set that can control the physical orientation and change in orientation to simulate many types of movement. This helps give sensation to the user, complementing the visual cues given on screen.

In this type of set up, all of the video and physical information is pre-processed, and so as the simulation runs, there are no issues concerning the speed of update of the physical set up, because the update speed is known before rendering.

However, in a real time system, latency and refresh rate are a serious issue. A screen can change, update, up to 120 times a second. However, an environmental set up will have very different refresh rates depending upon its' type. In order to give an accurate representation of a virtual environment at any stage, an interactive environment will require an interactive rate of update.

C.2 Personalised cues

These cues are given individually on a per user basis, and have a great deal of value in terms of communicating both discrete and continues effects, such as to communicate physicality of virtual objects, to effect feedback according to a response etc.

Individualised user policy

Each user can be treated according to their needs, the equipment and thus the experience tailored to cope with different people.

Exclusions

For a personalised set up, a number of types of exclusions must be present. Users may be required to be physically excluded from each other in order to use a piece of equipment, or they may need to be physically excluded from their environment in order to use the equipment. Either way, this exclusion has implications upon the user and their ability to use to experience physical information in a virtual environment.

C.3 Implementation options

Environmental options

There are many options for implementing an environmental Haptic feedback system, mostly derived from the entertainment industry. These include but are not limited to : omni-orientation platforms in the form of simulators, environmental effects as used in theme park rides such as wind forces / temperature changes. Whilst these are able to produce a series of effects that simulate physical environmental forces, they are not easily applied within a virtual environment in their current states, indeed there seem to be no common interfaces for such environmental Haptic systems outside of military application-based systems such as flight simulators, which are custom built.

Personal options

Where the environmental options are wide but not easily implemented, the research and home entertainment industries have produced a number of devices to be used as part of a personal system.

Research has yielded a few specialist devices, such as the omni-orientation Phantom, which both sends and receives orientation and position information, as well as providing force feedback. A selection of semi-immersive interaction and feedback devices is available from Immersion Ltd in the form of interactive gloves / exoskeletons that can either give tactile information via localised actuators in the tips of the fingers, sense from the joints the position of the fingers. In the CyberGrasp system, also available is the ability to provide force feedback and restriction of movement via a metal exoskeleton.

There are also a small number of devices that use sound, particularly very low resonance waves, to “augment” via vibration, such as using a chair or simply to enhance another device.

Haptic devices do not perform processing for themselves and need to be connected to a PC, either via USB, FireWire or special PCI card configuration. Software is needed to feed a signal to the device, and multiple users need multiple hardware and software configurations. This may deem a multi-user experience a multi-PC setup requirement, networked together, as opposed to a single PC responsible for the graphics and sound of an environmental model.

The cost of these systems varies quite considerably between these areas. Whilst the games industry based devices, mostly designed to be used on home machines and games consoles, go for as little as 20-30 pounds each, the immersion systems based Cybergrasp is around 68,000 pounds per setup.

Below is a breakdown of possible devices, rough estimates at cost, implementation aspects and possible usability issues:

Haptic devices are broken into two main areas :

(a) Augmented Human devices

These devices are assigned to a human in order to directly interact, either using total immersion (e.g. full body suit containing physical or electric sensors and/or actuators) or

partial (particular areas, such as arms / legs / restrictive head movements). These types of devices are the most sophisticated but also the most pricey, a good example being the CyberGrasp system by Immersion, which combines position and orientation sensing (input) with force feedback via an actuated exoskeleton.
These types of device are only as useful as their application, immersion within a virtual environment may not require the use of all the different systems, for example simple walkthroughs may not require a system such as the CyberGrasp except to handle simple objects, open door, shake hands with avatars etc. They could, however, be useful in rendering more general environmental effects, such as simulating the force felt when walking through water etc.

(b) External devices

These devices are often a metaphor for devices in the real world, and use generated electromagnetic forces as a way of simulating the effects that using the real device would generate. These often have very specific use in an environment, such as steering a vehicle using a force-feedback steering wheel or joystick, using a device such as the Phantom Omni as a pointing / force feedback device, or a simple game pad as a means of both inputting commands and feeling responses / events that have occurred.
The combination of external devices with sound can also be a powerful of haptics in the form of low level audio. There are devices which are concerned with adding force feedback to the users' seating to provide low level rumbling, as an environmental form of feedback.

Below is a breakdown of possible haptic devices, implementation aspects and possible usability issues:

- Game pads
  - Around 20 pounds or more
  - Vibration only for positive or event reinforcement
  - USB or Gameport requires WindowsXP, 2000, ME or special installation
  - Also allows user input, individual button and analogue stick input
  - Individual users can enjoy customized use.
  - Users may need training each time in use.
- Joysticks
  - Vibrations and force feedback restrictions movement and gives feeling of inertia
  - USB or Gameport requires WindowsXP, 2000, ME or special installation
  - Also allows user input, directional / orientation input
  - Individual users can enjoy customized use.
  - Users will not need training in use, as it is intuitive.
- Steering wheels
  - Vibrations and force feedback restrictions movement and gives feeling of inertia
  - USB or Gameport requires WindowsXP, 2000, ME or special installation
  - Also allows user input, wheel based turning and analogue input via a stick.
  - Individual users can enjoy customized use.
  - Users will not need training in use, as it is intuitive.
- Chairs
  - Range in price from around 1500.
  - Range of audio sub-woofer based vibration feedback devices, as well as one orientation device.
  - Connect to specialist hardware PCI card in PC.
  - Individual users can enjoy customized use.
  - Users will not need any training in use.
- CyberGrasp / CyberTouch systems (partial augmented humans)
  - Allow full immersive hand position sensing and orientation input
  - Allow feedback in terms of restrictive movement (CyberGrasp), fingertips actuators (CyberTouch)
  - Specialist hardware PC setup
  - Users will require careful training in use as equipment is fragile.
  - Individual users can enjoy customized use.

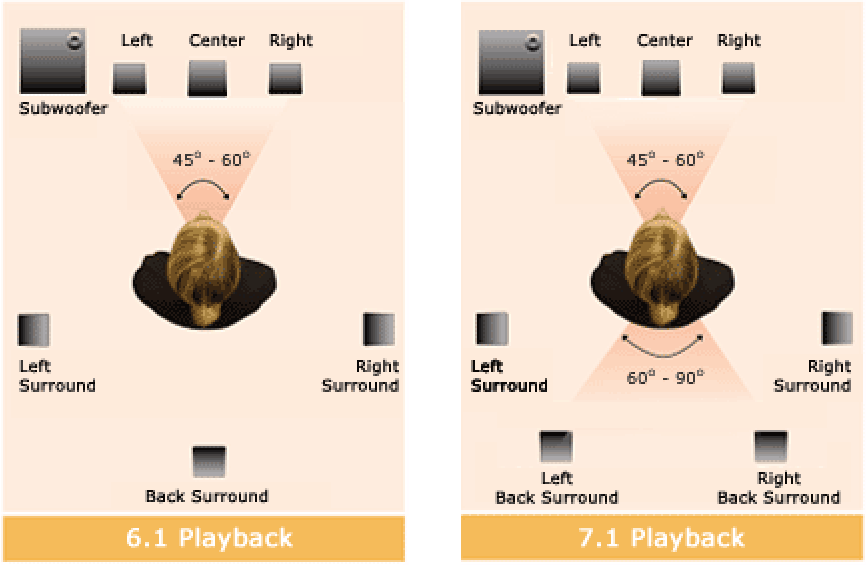

Figure 12

9.1.2. Input Devices

Input devices for virtual reality installations are an extension of traditional Human Computer Interaction devices with an emphasis on navigation, object handling, choice selection & dialogue interactions. In order to enhance the realism of any simulation, it is necessary to increase the similarity of the use of a particular input device to the action performed in the simulation. However this principle is often compromised because in many simulations an input device must fulfil several roles. The following subsections consider characteristics that govern VR input technology and how these characteristics affect currently available input technologies. Categorisation of such input devices is difficult on account of the large variation in methods of interaction and data they provide. Two existing taxonomies are described by Baecker & Buxton (1987) and Card, Mackinlay & Robertson (1990).

A. Input device characteristics

Modality of interaction

Modality of interaction governs which actions a user must take to interact with the input device. Examples of different modalities are hand motion, walking, speech & eye movement. Many tasks can be completed by a number of different modalities of interaction though some are more intuitive for users. One measure of the appropriateness of a modality to a particular form of input is referred to as the "Gulf of Execution" (Hutchins et al. 1986).

Multimodal input systems can increase a user's sense of immersion whilst allowing them to interact more rapidly with a system. A drawback of multimodal systems is that the additional complexity of such interactions can result in greater training periods.

Dimensionality

All input devices are governed by the number of independent channels that they can pass information from the user to the system. Each of these channels representing a dimension. Where input devices are used for navigation it is normal for as many as 6 dimensions to be required. Three are for linear movement and a further three for rotational movement. Within heritage VR environments it is quite likely that navigation in the additional dimension of time could be required.

Where input devices have less channels available than the number of dimensions they are required to interact with then it is still possible to use them provided that different dimensions can be interacted with in a sequential manner.

Discrete vs. continuous

Input can take the form of a continuous signal or a discrete value. More information can be passed between the user and system where a continuous channel is employed but this can be at the expense of harder control for the user.

Tracking vs. selection

Rather than having to make intentional signals to interact with a system through an input device, it is possible for a more naturalistic user experience where the users can have parts of their body or the whole of it tracked and their natural motions can be used as the input device. Tracking devices can result in very quick training periods for users.

Where selection input devices are employed the user is able to decide whether they wish to make a particular interaction with the system but this is not always possible where a tracking input system is used.

A potential problem for some tracking devices is the space requirement. In order to make movements that are possible within the system, it must clearly be possible to make such movements in reality. A constricted space for the user who is being tracked would not allow them to take full advantage of unencumbered space within the virtual environment.

Bandwidth

The time taken for the intentions of a user to be acted upon by the system can have significant effect on the immersive sensation experienced by a user. As this time increases, so a lag will grow between a users action and the effect the user will experience with this action. For the input device two factors affect this rate, the time for the user to make the physical action and the time for the input device to encode this action and transmit it to the application (Card, Mackinlay & Robertson 1990).

For VR simulations particularly this lag time needs to be as short as possible as the application handling the input will add further lag before feedback is created for the user. For most input devices interface circuitry rather than the sensors presents the greatest delay (Mulder 94).

Precision

The precision of any particular device will determine the appropriateness of the device for any particular task. Within VR simulations where the user has immediate feedback from input actions, high degrees of precision are not required as the user is able to easily adapt to the behaviour of the input device.

Reconfiguration

Due to a number of restrictions such as space and cost, it is likely that an input device is to take many roles within a virtual reality interaction, therefore it has to be as appropriate as possible for the different positions. For such input devices there is a trade off between having a complex device with simple separate functionality for each type of input against a simpler device with few input channels but each being multifunctional.

Ease of Use

On account of the wide variety of possible input devices, many will be unfamiliar to users, the first time they interact with a particular VR environment. Some input devices are more intuitive to use than others though this is very dependent on the context in which they are used in the VR simulation.

Size and shape of the input devices can influence the ease with which users can employ the input devices. Weight, position, sensitivity and requirements of mobility all can have significant effects on limiting or excluding users.

Privacy

Privacy of a user's interactions with a system can be influenced by the type of input device used. User privacy can affect some users' reticence to engage with an environment and also create difficulties on account of interference of overlapping input where two users are in close proximity.

Cost

The cost of input devices can vary quickly particularly where new technologies are being adopted. Given the enormous gulf of price differences between input devices cost acts as an important differentiator for products suitable for large installations verses professional tools. Cost is very significant where there are multiple users with each user having a separate input device. Further, in public installations where it is likely that input devices will experience heavy use, the cost of maintaining and servicing must be considered.

Durability

Input is an interface between the user and the computer system so by its very nature it must withstand interactions with the hardest user it will encounter. Within a museum and heritage environment this is the general public. Where an installation is to be supervised, devices are required to be less durable than environments which are unsupervised. Durability often will have an effect on sensitivity of an input device with more durable devices being less sensitive. Where an input device has separate detachable components then its durability will decrease with a higher risk of such components being lost. A further requirement for any input device is to withstand continuous cleansing where it is to be shared amongst many users.

B. Input device technologies

Input devices for VR systems have emerged from GUIs in desktop computer technology. These devices fall into two broad categories, trackers and selection input devices. Trackers allow a user to behave in a normal fashion as though they are interacting with a real version of the environment that is being simulated. The user's motions are then captured and this information is used to update the simulation to reflect how the environment would change relative to the user if they had undertaken a similar action in the real world. Intentional input devices allow a user to make specific actions which map to arbitrarily defined (though good design will dictate that they are intuitive) behaviours within the VR system.

Within these two broad categories the various input devices are then organised by modality of interaction. The most common being with users hands and fingers though body movement, eye movement and speech are also considered.

<u>Selection Input Devices</u>

(a) Game pads / joysticks

Joysticks are an intuitive interface for navigation and spatial control and have been adopted as a computer interface from the field of avionics. A single joystick offers continuous input in 2 dimensions and with some form of toggle switching this dimensionality can be increased indefinitely, though at the cost of only 2 dimensions being accessed simultaneously. The adoption of joysticks by the gaming industry in the form of game pads that typically integrate 2 joysticks and a number of discrete buttons into a single unit has ensured that the input device is both familiar, cheap and robust. The use of 2 joysticks simultaneously ensures that game pads have double the bandwidth of a single joystick, though this comes at the cost of new users taking longer to learn to use the device.

(b) Keyboard

The keyboard is very familiar for many users but difficult to use fluidly for those not experienced even with a short training period. Furthermore, the layout of keyboards varies from one language to another even where similar alphabets are shared so within a multinational context such as tourism complete familiarity will not be common. Furthermore it is difficult to convey navigation controls as there is not an intuitive correlation between the discrete keys and continuous spatial movement. The ability of a keyboard to transmit language means that this input device is most suited to natural language like interactions, issuing commands and responding to menu selections. The complexity of the keyboard as an input device could result in it being distracting to the user who might need to be concentrating on the keyboard rather than the screen. Keyboards are both cheap and in the case of membrane keyboards, robust.

(c) Mouse / trackpad

The mouse is one of the most familiar input devices for computer interaction, offering continuous input in 2 dimensions. Like the joystick, the mouse incorporates additional buttons that can be used for selection and toggling to adjust the dimensions that the mouse movement controls. Unlike the joystick, the mouse is not a fixed device thus reducing its robustness. Trackpads, which now replace mice on portable computers, offer a similar form of interaction but can be made more robust as they do not need to be moved.

(d) 3D mouse / spaceball

3D mice offer a similar form of interaction to the traditional mouse, but with additional dimensionality. These are a less familiar form of computer interaction than the traditional 2D mouse. An extension is the Space ball that allows the user to interact with an additional 3 dimensions by monitoring not only translation but also rotation. These dimensions map well to typical dimensions required to explore a 3D environment.

(e) Touch screen

Touch screens are computer monitors which register where on the monitor, in relation to the pixels that they display, they are touched by the user. Software is thus able to determine both discrete presses and movement made by the user. By adjusting the display according to the type of input required from the user, the touch screen can simultaneously simulate any number of discrete buttons (even a keyboard), trackpad like devices and context relevant buttons such as hypertext and graphical hyperlinks such as image maps.

An account of the possibility for direct correlation between user action and the point of interaction, touch screen input devices can be very intuitive. This is further enhanced by their ability to dynamically update the look of the point of interaction to provide any further information that helps with that interaction.

Touch screens offer the advantage of highly reconfigurable input devices that allow many types of input to be performed in a little space. This flexibility comes at a cost for additional development time. They are durable (typically 30-50 million presses at any single point).

(f) Voice recognition

Voice recognition allows users to interact with a system through both natural language by dialogue and the issuing of commands. This is achieved by the user talking into a PC microphone and the resulting signal is analysed by speech recognition software on the PC. Such software currently places a relatively expensive computational burden on the system and the cost of such an input device must incorporate both a dedicated PC and the voice recognition software.

Currently voice recognition software presents difficulty with a multilingual audience as systems not only are trained for a specific language but also work better for specific dialects and accents. Voice recognition systems also work better after a training period in which both the user and the system can adjust towards better interactions.

Where voice recognition is to be used for natural dialogue, current technology must limit this to specific domains in order to be able to handle the complexity of language understanding (Martin et al).

(g) Gesture recognition

Gesture recognition input devices offer a natural way for users to interact with a system even if most users would currently be unfamiliar with such an input device. Hauptmann notes that it is intuitive for all.

"there are no expert users for gesture communications. It is a channel that is equally accessible to all computer users" (Hauptmann 1989 p. 244)

Two forms of input device that exploit gesture recognition are input gloves (Krueger 1991) and hand trackers such as the Vectorix and Phantom (Sensable 2005). In the case of the glove a user places this over their hand and it monitors movements of both the hand and the fingers. This allows for a multidimensional input device as the hand is able to move with six degrees of freedom and each finger can have several degrees of freedom. With the hand trackers, the user can move a limb around in space and in the case of the Phantom haptic feedback is possible. Both approaches are easy to manipulate but training is required to link different motions with input channels that are not spatially correlated (MacKenzie 1995).

The use of a glove as an input device presents some difficulty from a user acceptability perspective. One issues is on account of a device that works for all user hand sizes and the other is with respect to hygiene maintenance (MacKenzie 1995).

Trackers

Tracking devices and wands allow natural motions of users and objects manipulated by users to be tracked moving through space. Three technologies dominate this domain, inertial, magnetic and visual tracking. Each of these approaches presents unique difficulties. Such input devices also allow for directed input through gesture recognition systems.

(a) Magnetic Trackers

Magnetic tracking devices require a user to wear or carry a transmitter whose signal is picked up by a number of sensors within close proximity of the user and is then used to determine both the position of the transmitter and its orientation. If the transmitter is tightly coupled with the user then this also provides the same information about the user. For magnetic trackers there are restrictions in movement range of the users. Magnetic trackers such as Ascensions 'Flock of Birds' (Ascension 2004) offer ranges between 1meter and 3meters and the user must support a cube with dimensions of 20 cm.

Three systems are currently employed relying respectively on , DC electromagnetic (EM) pulse, AC EM field strength, and AC EM field. The DC approach suffers from interference from magnetic fields such as the Earth's whereas the AC approaches place restrictions on environments to not include additional ferro-magnetic materials (Mulder 94).

(b) Inertial trackers

Inertial tracking systems can be quite robust and require relatively little additional computational power to provide information about their positioning. With the potential for wireless integration, they can also be quite unobtrusive. A problem with such systems is that they need to be recalibrated frequently as errors can act accumulatively.

(c) Visual trackers

Visual tracking systems work by using a series of cameras to monitor user movements. Thus a user is free to move around an environment and their movements are automatically monitored to update the environment. This can be achieved through the identification of body shapes or more easily where users have markers placed on specific parts of their bodies. The placing of such markers can present difficulties both from the perspective of privacy and accuracy.

Visual tracking systems tend to be quite expensive on account of the high speed cameras and the extensive computational power required to analyse scenes. Furthermore, the large amount of computer power required can mean that the response times of the input device might be larger than is acceptable for some VR applications.

The arena in which visual tracking works must be restricted so that the user can never be too far away from cameras in order that the cameras can resolve the actions of the user and the lighting must be controlled so that it is quite constant throughout (Mulder 1994).

A particular problem to affect visual tracking systems is occlusion. More cameras used by the system can help to partially alleviate this problem (at the expense of additional computational burden) but it will still exist particularly where there are multiple users within a single arena but also with smaller body parts such as hands with a single user.

(d) GPS trackers

GPS offers low accuracy at present with precision limited to approx 30cm. A further problem with this technology is that signals are very weak so it will not work indoors. An advantage is that there is no limit to the area it will cover (Mulder 1994).

(e) Eye trackers

In visually-mediated applications then eye trackers can be a powerful input device (Duchowski 2002). Typically eye trackers provide an accuracy of between 0.5 – 2.5 degrees with a response time of 100 milliseconds (EL-MAR 2002). As well as monitoring direction of gaze, trackers can also be configured to recognise certain combinations of blinking actions to activate switching behaviour.

Trackers are either mounted on a headset, or separately from the user. Where the tracker is mounted separately, the user is restricted to motion within a small area in front of the tracker, typically 10-20cm motion in any direction (Duchowski 2002). Where the tracker is mounted via a headset, the headset includes a camera (or two where stereo eye tracking is employed) which is not very robust. In order to establish the direction of gaze, the user must adopt a fixed position, normally with some chin rest, or their head position and orientation must also be tracked and integrated with direction of gaze (Huang et al. 2004).

The difference in cost between the head mounted systems and separated systems can vary by over an order of magnitude.

(f) Tread mill / walking detectors

Both single and multidimensional treadmills have been developed to be used for VR input. Examples of walking detectors are the “Fantastic Phantom Slipper”, an infrared –emitting pair of shoes that are tracked and “Joyfoot” which uses acceleration sensors (Barrera et al. 2003). Whilst such devices can offer more immersive VR experiences, they can also act as an encumbrance as navigation through a large virtual world might be achieved too slowly through walking or running. Furthermore, such an input device can act as a total obstacle to users with restricted mobility.

(g) Biometric measures…

Further input devices where the user has less control are biometric monitors. Examples of such devices are the heart rate monitor and skin surface conductivity meter which could help determine a user’s emotional response to different parts of a system. Work has also been pioneered to use EEG as an additional source of input (Freidman et al. 2000). Caution must be taken with regard to user privacy with such intrusive devices.

C. Input device conclusions

The suitability of a variety of input devices has been considered for Virtual Reality systems within the cultural heritage domain where the main tasks are navigation, object handling, language interaction and choice selection. Input devices offer a way for users to interact with virtual environments and where the actions required to operate input devices correspond to similar actions performed within the virtual environment then an intuitive understanding of the device can be readily achieved as well as a more immersive experience. A joystick does not provide an easy interface for conversation and speech input is not the easiest method for navigation. However, other considerations can outweigh these benefits. For example, by tracking walking behaviour a user could navigate in a very intuitive manner through a virtual environment, but space restrictions as well as cost might dictate against this as an input device and an alternative of using a joystick could be more successful.

The input devices of a system must provide a means for users to explore all aspects of interaction of the system. This means that a device must have sufficient bandwidth for all degrees of freedom within a system to be manipulated. This bandwidth is not just between the input device and the system but also between the user and the device. Clearly additional bandwidth can be achieved where several devices are employed simultaneously such as the two joysticks and many individual buttons on a gamepad, as long as they employ different modalities of input from the user, but this approach can get harder for users to learn.

By employing an input device that is reconfigurable then it can adapt to be used for many types of input rather than relying on the user having to switch between devices or use a single device often in a non optimal way to perform many types of input.

9.1.2. Haptics in cultural heritage

Haptic technologies involve the sense of touch in computing (from the Greek haptikos meaning "able to touch"). Recent developments in Virtual Reality currently allow computer users to use their sense of touch to feel virtual objects. State-of-the-art haptic (or force-feedback) devices allow users to feel and touch virtual objects with a high degree of realism.

An artefact's surface properties can be modelled so that someone using a haptic device could feel it as a solid, three-dimensional object with different textures, hardness or softness. Haptics is very much a device-driven technology; these devices include pens, gloves, joysticks/ joypads and force-feedback mice. Haptics involve input/output peripherals which range from rather cheap to very expensive (e.g. a top-of-the-range SensAble[74] PHANToM haptic interface with six degrees of freedom *(right)*.

These devices enable users to manipulate virtual representations of objects and artifacts from remote workstations. Web-delivered haptic exploration is beginning to mature and collaborative haptic sessions which take network delays (or 'latency') into account are now emerging (Digicult TWR 3).

There is an active community of engineers and computer scientists involved in the field of haptics. The Haptics Community Website[75] is intended to be an online resource for researchers in the haptic display community, while the International Society for Haptics[76] is a professional society bringing together researchers interested in haptics. Haptics-e[77] is an e-journal containing various information on haptics research and use. Several international symposia and conferences are also devoted to haptic research, e.g. the World Haptics Conference[78] and the Symposium on Haptic Interfaces[79].

[74] http://www.sensable.com/

[75] http://haptic.mech.nwu.edu/

[76] http://www.isfh.org/

[77] http://www.haptics-e.org/

[78] http://www.worldhaptics.com/

[79] http://www.hapticssymposium.org/

Some recent European research projects concentrated on haptic technologies as well. The GRAB project[80] for example has allowed the development of a new Haptic & Audio Virtual Environment to allow blind and visually impaired persons to have access to the three-dimensional graphic computer world through the senses of touch and hearing. The new GRAB system allows its user to interact and explore 3D virtual objects both with the thumb and the index fingertips and with both index fingertips while moving the hands in a desktop workspace.

Another EU project focusing on innovative haptics technology was called Tacitus[81]. The principal aim of the project was to investigate the development of three-dimensional haptic and multi-sensory computer applications for creative processes in applied arts and design.

Haptics in cultural heritage

Brewster (2003) describes the background to haptics, some of the possibilities of haptic technology and how they might be applied to cultural applications. Haptics are a potential solution to overcome a common restriction for museum visitors: the tactual exploration of works of art and archaeological objects. Tactual exploration could represent an important means for fully appreciating the features of this kind objects which are not readily available or which are too fragile or too valuable to be touched (Bergamesco et al 2002a).

Safely and effectively exhibiting such objects is always a challenge. On one hand, the museum wants to make the object as visible and accessible as possible to the public, while on the other, its access must necessarily also be restricted due to security factors and the object's fragility. Access to this kind of collections thus can be enhanced through haptic interfaces, reducing time or distance constraints. Some examples include: making very fragile objects available to scholars, allowing visitors who live far from museums to feel objects at a distance, letting visually-impaired and blind people feel exhibits that are normally behind glass, and allowing museums to show off a range of artefacts that are currently in storage due to a lack of space.

As the advantages outlined above, mostly concern museum professionals, the application of haptics in cultural heritage is more developed in the museum sector. Haptic technology is already being used in museums, but on a small scale in very specialized situations. One such is the University of Southern California's Interactive Art Museum[82].

In 2003 a workshop at the Museums and the Web conference[83] was dedicated to the theme of haptic technologies. In December 2004, an entire conference was organized by the Heritage Studies Research Group, Institute of Archaeology, University College Londen. (The magic

[80] http://www.grab-eu.com/

[81] http://www.eca.ac.uk/tacitus/

[82] http://digimuse.usc.edu

[83] http://www.archimuse.com/mw2003/abstracts/prg_200000719.html

touch. Touching and Handling in a Cultural Heritage Context'[84]). This clearly proves the interest of museum professionals in the field of haptics.

In contrast to the clear advantages outlined above, the application of haptics in cultural heritage (and also in other sectors) is still hindered by several obstacles. At present the costs for applying this technology is still too high for the average CH institution. Not only are precise haptic devices quite expensive, also the 3D scanning and modeling of the artifacts involves high costs. Museum professionals might also raise hygienic concerns, as with any kind of hardware installed in museums and used by large numbers of visitors.

In other fields of cultural heritage, haptics still have to prove their potential. The statement by Ferko et al. (2003) that "unlike audio and video, haptic devices are quite rare and therefore we see no relevance of haptics for Virtual Archaeology now, except the possible Braille interface," is a clear indication.

Only few research projects have dealt with haptics in cultural heritage during the last years. The Museum of Pure Form is probably the most interesting and promising example. Another project, CREATE, also tries to improve the interactivity in virtual worlds

<u>The Museum of Pure Form</u>

The Museum of Pure Form[85] aimed to explore new paradigms of interaction with sculptural pieces of art and create a virtual gallery featuring three-dimensional digitised sculptures from European and worldwide museums with which visitors could interact with touch and sight. Users can perceive subtle tactile stimuli which equate to the contact of their hands with the digital copy of a real object.

The Museum of Pure Form system is composed by a video subsystem of 3D projection and a haptic subsystem for force feed-back. The haptic subsystem reproduces the contact forces generated by the interaction with the digital model through a haptic interface, so that the user may perceive a replica of the forces generated during the contact on his/her hand, together with the virtual representation of the movements of his/her hand along the surface of the digital model.

[84] http://www.ucl.ac.uk/archaeology/events/conferences/magic-touch/index.htm

[85] http://www.pureform.org/

Two different realizations of "Museum of Pure Form" systems were developed at PERCRO[86], Pisa, Italy. In one realization "The Museum of Pure Form" is conceived as a system laced inside several museums and art galleries around Europe and made available to people visiting such institutions. In the second realization the system will be laced and tested inside a CAVE environment. Considerations on technological aspects concerning the implementation of a virtual environment endowed with haptic feedback for cultural heritage applications are reported in Bergamasco et al. (2002b).

The CREATE Project: Mixed Reality for Design, Education, and Cultural Heritage with a Constructivist Approach

The global scope of the CREATE[87] project is to develop a mixed-reality framework that enables highly interactive real-time construction and manipulation of photo-realistic, virtual worlds based on real data sources. Compared to previous research and design in virtual worlds, the CREATE project uses a high degree of interactivity, and includes provision for other senses (haptics and sound). The applications developed in CREATE are designed to run on different platforms, and the targeted running systems are SGI and PC driven, with immersive stereo-displays such as a workbench, a ReaCTor™ (CAVE-like environment), and a wide projection screen (Loscos et al. 2003, Roussou & Drettakis 2003.).

Hong & Woo (2004) proposed the I2-NEXT, which enables users to interact with virtual objects by haptics and tangible objects in immersive networked virtual environment. The proposed system provides users with haptic interaction to experience force feedback when they interact with virtual objects. It also supports large scale stereoscopic display through clustering technique. As to show the effectiveness of the proposed system, Hong & Woo have been developing an application for digital heritage reconstruction.

[86] http://www.percro.org

[87] http://www.cs.ucl.ac.uk/create

9.2. Future directions and EPOCH's contribution

One of the important activities in the first year of EPOCH's activities was the realization of eight showcases. One of them also included haptic interfaces, in order to develop a multimodal interface for the safe presentation of valuable objects.

The showcase[88] is based upon a replica of the object which serves as the interface to explore the object. Through the use of an orientation sensor integrated in the replica, the object can be visualised on a computer screen in precise coordination with the angle it is held or rotated by the visitor/user. In this way, the user feels the shape and details of the object and sees the virtual representation of the object in the screen, behaving exactly the same way as the replica. By adding touch sensors to the surface of the replica in significant areas of interest, the user can explore the meaning of the object. Touching a selected feature on the surface of the replica brings up a story on the screen that explains some facet of the meaning and history of the object. For example, if the object bears an inscription, the user can learn what it means, and what message lies behind it, by simply touching the text. This tactile interface allows visitors to experience and explore the object in an exciting and innovative way — that would certainly not be possible with the original artifact.

The application is based upon the ARCO[89] software of the University of Sussex, and a concept developed by the Ename Center. The replica is made through stereo lithography of a 3D model of the object, obtained by laser scanning. EPFL and the Ename Center create the 3D stories with virtual reconstructions of the Ename abbey and its inhabitants. The visualisation of the object is done through the most recent 3D flat screen visualisation techniques.

Besides this showcase, several partners of EPOCH do have extensive experience with haptic technologies (e.g. Un. of East Anglia and ETH Zurich, where several types of haptic displays are installed; ETH organised EuroHaptics 2004, after co-founding the conference, and are partner in the 6FP project TouchHapsys).

Expertise on haptics is available in the following partners of the EPOCH network:

- Brunel University
- University of Hull
- University of Surrey
- ETH Zürich
- University Patras
- ...

[88] http://www.EPOCH-net.org

[89] http://www.arco-web.org

Large-scale and scarce facilities are available at different partners’ sites:

- Cave systems (Braunschweig, Fraunhofer, ETH)
- Dome (Bonn, Leuven)
- VR room (ETH, Imagination, Miralab)
- Optical lab (Valencia)
- Multimedia, visualization, usability labs (Imagination, 4site, Brighton, Interactive Institute…)
- Haptic displays (ETH, liaising with Koc Un.)
- …

CONCLUSION

Haptic technologies are not expected to become available for the vast majority of cultural heritage institutions in the coming years. There are several reasons for this, which came out of our discussions with (mainly museum) partners: 1) the technology still is quite expensive, 2) it is rather fragile and may not survive visits of, say, a school, and 3) the technology serves one visitor at the time, which raises issues of throughput, esp. for high-profile exhibitions with large visitor numbers (but exactly the type of event where it could still be financially viable). These conclusions are consistent with a summary that DigiCULT (5FP) has published about ICT for cultural institutions at ICHIM 04: ` haptics… these technologies will most likely remain beyond the reach of small- and most medium-sized institutions ‘ (also in the longer term of 6 years and beyond).

However, the new CyberWalk project will make use of the NoE’s 3D site models, taking Pompeii as a starting point. CyberWalk will enable quasi-natural, unconstrained, and omni-directional walking in virtual worlds.

9.3. References

Ascension (2004) *Flock of Birds specification* http://www.ascension-tech.com/products/flockofbirds.php

Baecker, R. M., & Buxton, W. (Eds.), (1987). *Readings in human-computer interaction: A multidisciplinary approach.* Los Altos, CA Morgan Kaufmann, 357-365

Barrera, S., Romanos, H. Takahashi, S. and Nakajima, M., (2003) Real Time Detection Interface for Walking in CAVE, *Proceedings Computer Graphics International 2003*: 105-111

Belleman Robert, Stolk Bram, de Vries Raymond (2001) Immersive Virtual Reality on commodity hardware, Proceedings of the 7th annual conference of the Advanced School for Computing and Imaging, pages 297--304, Heijen, the Netherlands, May 30-June 1 2001

Bergamasco M., Brogni A., Frisoli A., Salvini F., Vignoni M. 2002. Tactual Exploration in Cultural Heritage. In *Proceedings of the XIV Computer-aided Egyptology Round Table*.

Bergamasco, M., A. Frisoli & F. Barbagli. 2002. Haptics Technologies and Cultural Heritage Applications. *Computer Animation* 2002: 25.

Bohannon William, (2004) Large-screen displays, Part One, 2004, Presentation

Brewster, S.A.(2001). The Impact of Haptic Touching Technology on Cultural Applications In *Proceedings of EVA2001*. (Glasgow, UK): 1-14.

Cambridge Display Technology (2005) http://www.cdtltd.co.uk/technology

Card, S.K., Mackinlay, J.D. and Robertson, G.G, 1990 The Design Space of Input Devices *Proceedings of the CHI '90 Conference on Human Factors in Computing Systems*,. New York: ACM.

DeBoer Clint (2004) Display Technologies Guide: D-ILA Projection, http://www.audioholics.com/techtips/specsformats/displays_D-ILA_projectors6.html

DeBoer Clint (2004) Organic LED Displays (OLEDs) - The Next Trend?, http://www.audioholics.com/techtips/specsformats/organicOLEDsdisplays.php

DigiCULT links : haptics, telepresence, robotics: http://www.digicult.info/textonly/links.php?t=19

DigiCULT TechWatch Report 3. 2004. http://www.digicult.info/pages/techwatch.php

Duchowski, A., 2002 *Eye Tracking Methodology: Theory & Practice* London, UK: Springer-Verlag (London), Inc.

Ferko, A. M. Grabner, G. Schwann, M. Sormann & K. Schindler. 2003. Navigation Functionality for Virtual Archaeology. In Proceedings of the 11th International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision. (http://wscg.zcu.cz/wscg2003/Papers_2003/J97.pdf)

Geser, G. & A. Mulrenin. 2004. Are small heritage institutions ready for e-culture? *Proceedings of ICHIM Berlin 04. Digital Culture and Heritage* (CD-ROM), September 2004.

Hauptmann, A. G. (1989). Speech and gestures for graphic image manipulation. *Proceedings of the CHI '89 Conference on Human Factors in Computing Systems*, 241-245. New York: ACM.

Hong & Woo. 2004. I2-NEXT: Digital Heritage Expo. ICAT04. http://uvr.kjist.ac.kr/papers%5CICAT04-Abstract_I2-NEXT.pdf

Howard, Webster E. (2004). Better Displays with Organic Films. Scientific American, p. 76

Huang, H., Allison, R. and Jenkin, M. 2004 Combined Head – Eye Tracking for Immersive Virtual Reality *The proceedings of the ICAT 2004 Conference*

Hutchins, E.L., Hollan, J.D. and Norman, D.A., (1986) “Direct Manipulation Interfaces” in *User Centered Systems Design: New Perspectives on Human-computer Interaction*, ed by Norman, D.A. and Draper, S.W. pp.87-124, Lawrence Erlbaum, Hillsdale, NJ.

Inition (2005) http://www.inition.co.uk/inition/guide_stereovis_autostereo.htm

IST, Displays and large area electronics, http://www.cordis.lu/ist/ka4/vision/irg_displays_and_large_area_electronics_2705.htm, last accessed on February 2005

Jansson, G., M. Bergamasco & A. Frisoli. 2003. A new option for the visually impaired to experience 3D art at museums: manual exploration of virtual copies. *Visual Impairment Research*, Vol. 5 No. 1, pp. 1–12

Kjeldskov, J. (2001). Interaction: Full and partial Immersive Virtual Reality Displays. *Proceedings of IRIS24*, Bergen, Norway.

Krueger, M. W. (1991). *Artificial reality II*. Reading, MA: Addison-Wesley

Loscos C. et al. 2003. The CREATE Project: Mixed Reality for Design, Education, and Cultural Heritage with a Constructivist Approach. In *Proceedings of ISMAR 03, The Second IEEE and ACM International Symposium on Mixed and Augmented Reality,* The National Center of Sciences, Tokyo, Japan, Oct. 7 - Oct. 10, 2003.

MacKenzie, I. S. (1995). Input devices and interaction techniques for advanced computing. In W. Barfield, & T. A. Furness III (Eds.), *Virtual environments and advanced interface design*, pp. 437-470. Oxford, UK: Oxford University Press.

Majumder Aditi, Stevens Rick (2005) Perceptual photometric seamlessness in projection-based tiled displays, ACM Transactions on Graphics (TOG) Volume 24 , Issue 1 (January 2005)

Manual of EL-MAR VISION2000, “Video Eye-Tracking System”, Version 5.0 June 2002

Martin, J-C., Buisine, S. & Abrilian, S. 2004 *2D Gestural and Multimaodal Behavior of Users Interacting with Embodied Agents* Nice Project http://www.niceproject.com

Mulder, A. (1994) *Human movement tracking technology* Hand Centered Studies of Human Movement Project Technical Report 94-1 Simon Fraser University, Canada

Musgrave Garry (2001) Very Large-Screen Video Displays, Conceptron Associates

Musgrave Garry (2001), Choosing a Video Projector, Conceptron Associates

NEC/Mitsubishi, (2005) Monitor Technology Guide, http://www.necmitsubishi.com/support/css/monitortechguide/index03.htm#resolution

Qinetiq (2005) http://www.qinetiq.com/home

Opticality Corporation (2005) http://www.opticalitycorporation.com/technology/index.html

Sensable (2005) http://www.sensable.com/products/phantom_ghost/phantom.asp

Soneira Raymond (2005) Display Technology Shoot-Out: Part IV, ExtremeTech Headlines

Sharp (2005) http://sharp-world.com/sc/library/lcd_e/s2_6_3e.htm

Pabst Thomas (2000), 3D Benchmarking - Understanding Frame Rate Scores, http://www20.graphics.tomshardware.com/graphic/20000704/index.html

Parkhurst, D., and Niebiur, E. 2002. Variable resolution displays: a theoretical, practical and behavioral evaluation. Human Factors 44, 4, 611--29.

Roussou and Drettakis. 2003. Photorealism and Non-Photorealism in Virtual Heritage Representation. In *Proceedings of VAST 2003*, Brighton, United Kingdom, November 2003..

Thon Harald and Weinand Lars (2004) GeForce Go 6800 Reaches For Desktop Graphics Performance, http://www.tomshardware.com/mobile/20041108/index.html

Websites

http://www.dolby.com/consumer/home_entertainment/roomlayout.html

http://www.headphone.com/layout.php

http://www.overclockers.co.uk/acatalog/Online_Catalogue_Sound_Cards_11.html

http://www.overclockers.co.uk/acatalog/Online_Catalogue_Speakers_22.html

http://www.homerecording.com/sound_card_basics.html

http://reviews.cnet.com/Graphic_sound_cards/2001-3002_7-0.html

http://www.immersion.com/3d/

http://www.hitl.washington.edu/scivw/faq.html

http://haptic.mech.northwestern.edu/intro/gallery/

10. Virtual humans and other avatars

10.1. Systems, tools and technologies

An ‘avatar’ (meaning ‘descent in Sanskrit) is a representation of a human user in a virtual environment. The term is sometimes used incorrectly for agents or bots, which are representations of computer processes or programs, and may be represented in the form of a human or some other creature.

DigiCULT’s Technology Watch Report 2 (2004) provides an extensive introduction to avatar technology. In the twenty years or so that avatars have existed, they have also attracted a strong interest with their ability to enforce the social components of using of a computer system Avatars continue to be created to inhabit virtual worlds. The uses to which they have been put in recent years have become increasingly varied.

Possible cultural heritage sector applications include (DigiCULT 2004):

- Tour guides in virtual exhibitions;
- Presentation of performance heritage with the reconstruction of traditional dances and theatrical performances;
- The representation of visitors in online spaces where users benefit from a sense of human presence;
- In conjunction with haptic interface and CAVE technology the tactile exploration of models of objects which for conservation reasons visitors cannot handle;
- As mechanisms to enable the study of users of historical spaces and human interaction.

The benefits of including virtual humans into cultural heritage reconstructions are twofold: the realism of architectural models is increased by populating them; and, as well, it allows to preserve the intangible heritage describing how people in historical times behaved. Gaitatzes et al. (2004) have been developing some virtual reality applications focusing on the Ancient Olympic Games. In their reconstructions much effort has been made to recreate the feeling of the games and help the user/spectator be an interacting part of the edutainment activity.

Avatars played an important role to enhance the level of empathy and interactivity. Similar avatar athletes were included in the ARCHEOGUIDE system for users situated inside the stadium. The competing avatar athletes are rendered on the live camera image of the stadium. Realism is enhanced by the individual behaviour of each athlete and the fact that the user is nog isolated but can, instead, see other visitors present in his field of view (Vlahakis et al. 2001).

Another important advantage of avatars and other human agents is that they provide help for disabled users (e.g. by means of sign language for visitors with hearing problems, or talking heads reading the display texts for the visual impaired). To illustrate the potential of avatars for disabled users: in May 2005 Deaf Connexions[90] announced the launch of their new website, which will includes avatars giving information in sign language.

In the mean time a lot of commercial software has become available to model avatars. A selection of relevant avatar modeling software has been included in EPOCH's common infrastructure[91] , while also a document with reviews of several solutions is available[92]

The first avatars were already developed in the eighties. One of the first examples of avatars in cultural heritage was probably IBM's pioneering reconstruction of Cluny Abbey, already had a virtual guide or avatar, in the form of a mediaeval nun, who took one around the virtual reality model of the famous church (Veltman 1998).

However, only in the last years, avatar technology has been frequently applied for cultural heritage purposes and real-time virtual characters are becoming inextricably integrated strands of the new emerging digital cultural visualization fabric. Together with the advent of recent powerful, low-cost consumer graphics hardware accelerators and cinematic real-time rendering, they contribute to the new compelling forms of representation and reconstruction of tangible and intangible heritage artifacts acting as a pervasive part of global culture.

Research on real-time virtual characters includes several issues: hair simulation, cloth animation, virtual crowd simulation, artificial life methods for behavioral animation of virtual characters, VR/AR platforms…

The rendering of avatars has improved dramatically in parallel with improvements in graphical and other related technologies. Three-dimensional avatars are now the norm. Some avatars may transmit the actual voice of the person they present, and text-to-speech (TTS) technologies have made it possible for agents to 'speak'. Avatars are not static, nor need their appearance be fixed. Users often design how they look, what they wear, and how they behave. Also clothing is very important: Koutsonanos et al. (2004), for example, developed an efficient framework for the interactive simulation and editing of clothing over avatars.

[90] http://www.deafconnexions.org

[91] http://www.student.kuleuven.be/~m9605958/tools/epochmatrix.html#7E

[92] http://www.student.kuleuven.be/~m9605958/tools/avmodsoft.pdf

Through the implementation of avatars in the, the user can imagine properly the size of the objects, as well as the context of the scene, the temporal orientation can be facilitated, the historical information augmented and a visual impression of history is given. The main advantage of an interactive 3D avatar is the generated illusion to the user of the existence of a person to interact with as if it was a real person. This includes a radical change in order to gain fluency in the relation with the computer system. There are different levels of complexity concerning the avatars. If the 3D virtual character has the real-time speech capacity, the impression of dealing with a real character is improved (Leanizbarutia et al. 2003).

The use of an avatar to serve as personal guide and storyteller for each visitor can greatly enhance the effectiveness of an interactive storytelling system. Interpretation specialists have long commented on the highly impersonal nature of the "disembodied" narrator’s voice. The use of an avatar (or multiple avatars) can provide a distinct point of view, attitude, and physical appearance to enhance effect of the interactively created story on the particular visitor or visitor group. Another EPOCH showcase, concerned with Avatar-Based Interactive Storytelling, also yielded promising results, bringing the history of ancient Ename back to life through avatar-monks.

EPOCH also realizes that speech is required to make interaction as natural and engaging as possible. The avatar must be able to respond to questions from a range of users with replies in a number of languages. Tools are required that allow new speaking avatars to be quickly developed. To this end, mouth shapes for all the speech sounds are created by mixing six base morphs and applying them to the reference face. Realistic visual speech can be generated using this limited set of morphs. These concerns lay at the basis of a showcase about ‘multilingual avatars’, developed last year. Avatars have been added to the 3D city model of Wolfenbüttel. They can talk to the visitor and will be combined with a background crowd of city residents so that the model comes to life with an interactive information guide.

A toolkit for speaking avatars has been developed by Wilkinson et al. (2004) as part of a larger system for building virtual models of cultural heritage sites. The toolkit is aimed at non-technical people who have expert knowledge of a particular site. The avatars must make efficient use of system resources, be simple to develop and look as realistic as possible. The toolkit aims to provide a straightforward route to providing accurate mouth movements for a speaking and gesturing avatar that will act as a virtual guide for visitors to virtual environments for cultural heritage sites.

The rendering of virtual crowds is another intersting issue. In the ERATO project an interactive real-time scenario of a virtual audience in an ancient Roman odeon in Aphrodisias was created (de Heras Ciechomski et al. 2004). Based on historical sources, both the building and the people were reconstructed. Inhabited virtual heritage applications require careful balancing of computational resources between the visualization of environment and the visualization of people. De Heras describes several techniques that allow achieving high visual quality for a large number of virtual humans rendered together with a complex architectural model while still keeping interactive framerates. Moreover, he proposes a comprehensive framework for authoring of crowd scenarios. More results of the ERATo project appear in Thalmann et al (2004) and de Heras Ciechomski (2005a, 2005b).

Although graphical human modeling has been a long sought subject in computer graphics, when it comes to dealing with real-time applications, it raises a number of unique requirements for both computer graphics artists and researchers. The LIFEPLUS[93] project successfully addressed this issue. For the first time, one is now able to run a combination of software processes to create walking, talking people with believable clothing, skin and hair in real-time (Magnenat-Thalmann 2004). The platform VHD++ features integrated real-time virtual character simulation technologies. This key innovation allows the plug-and-play of different heterogeneous technologies such as: real-time character rendering in AR, real-time camera tracking, facial simulation and speech, body animation with skinning, 3D sound, cloth simulation and behavioral scripting of actions. In future research researchers at the MiraLab would like to give a better sense of presence by adding automatic behaviour of virtual characters when real people enter the virtual scene. They would like to develop research on interchanging natural consciousness between real and virtual, and between virtual characters.

Interactive tour-guide robots were developed in the TOURBOT[94] project. These robotic avatars are able to provide individual access to museums' exhibits and cultural heritage over the Internet. TOURBOT operates as the user's avatar in the museum by accepting command over the web that direct it to move in its workspace and visit specific exhibits; besides, TOURBOT can also act as a flexible, on-site museum-guide (Giannoulis et al. 2001). Also, the DHX[95] project aimed at providing ways of sharing and communicating heritage contents between museums worldwide with VR representations, interaction paradigms and remotely controlled avatars.

[93] http://www.miralab.unige.ch/subpages/lifeplus/

[94] http://www.ics.forth.gr/tourbot/

[95] http://www.eurasian-dhx.org/

10.2. Future directions and EPOCH's contribution

Different partners in the EPOCH network are currently doing research in the field of avatar technologies. Here we discuss a framework for infrastructure work in the area of virtual humans.

10.2.1. Proposed Infrastructure Framework

Standards

EPOCH will continue to gather information about standards for representing virtual human characters and for specifying animations for existing characters. Most of the de facto standards relate to tools described under small scale tools below and can suffer from the lack of open documentation associated with proprietary formats. MPEG-4 FBA provides a standard that is non-proprietary but is not in wide-scale use. Although it is desirable to be able to specify animations that can be used with a range of characters, most formats are avatar-specific.

Proposal

EPOCH will provide guidance for cultural heritage organisations on the practical standards to be specified when commissioning virtual human characters and animations for multimedia presentations.

EPOCH recommendations will be based on the availability of transformational tools to enable flexible reuse of assets conforming to the recommended standards. Such tools will assist in transfer of characters between tools, and retargeting of animations between tools and characters.

EPOCH recommendations will be based on the support of such standards from commercial and academic suppliers, both within the EPOCH Network and in the wider market.

Repository

It will be valuable to have a repository of virtual human characters and animations to support developments in EPOCH. Such examples will be valuable to test and demonstrate the ability to transform assets for use in a range of formats. A repository would provide test data for inclusion in prototype multimedia presentations. Assets in such a repository might be licensed for direct inclusion in complete presentations.

Proposal

EPOCH will support the chosen standards for virtual human characters and animations by collecting an open repository that can be used for non-commercial purposes in research and prototyping. The repository will illustrate the effect of applying appropriate transformation tools to convert assets from one format to another.

Small Scale Tools

This section focusses on the use of tools for creating and developing virtual human characters and animation sequences. Many such tools can also be used to create complete virtual environments, but in this section the concern is with the ability to develop assets that may be used with a range of tools. Work will be needed to verify the flexibility of the tools in use and the amount of manual intervention needed to reuse assets.

Proposal

EPOCH will identify a set of tools that can be used successfully with the standards identified for virtual human characters and animations. EPOCH will identify or aim to develop tools for transforming assets from such tools into standard formats. In the light of current usage, the following will be considered carefully: 3ds Max with Character Studio (Discreet), Maya (Alias), Motion Builder and FBX (Alias, previously Kaydara), Poser (Curious Labs) and DAZ3D. Tools from EPOCH partners, especially UNIGE, EFPL, and UEA will also be considered.

Integrated Tools

Complete multimedia applications for cultural heritage presentations can be built entirely within the framework of some of the tools mentioned above or by importing assets into such frameworks. Virtual humans will not be the main focus of cultural heritage applications but will be integrated into complex scenes built from 3D models either of existing objects and sites, or of reconstructions based on expert opinions. It will often be important to trade speed of rendering for fine detail and total realism so that real-time interaction is possible using inexpensive hardware. Work by EPOCH partners will have a strong role to play.

Proposal

A survey and checklist of major commercial tools will be provided based on ability to build presentations from EPOCH standard components. Service companies will be asked to state which formats they can handle.

In line with the proposals in the subarea of 3D visualisation, it is expected that OpenSG frameworks developed by EPOCH partners (TU-BS and UEA) will provide good support for virtual humans to be integrated into scenes, both as 3D geometric avatars and simpler representations for crowd members.

For control of virtual human behaviour the proposed VHD++ framework (MIRALab-UNIGE and VRLab-EFPL) is expected to provide support for standard virtual humans and animations. Integration between VHD++ and the EPOCH OpenSG framework will be explored.

Platforms

EPOCH needs to keep in mind a number of different platforms for deploying multimedia presentations. The issues here are in common with the 3D visualisation subarea. For most applications, it is essential that they can be presented in real-time on reasonably priced commodity hardware. For some advanced presentations it will be important to support high-end processing systems and novel displays. Some high-quality presentations will require offline processing, but will lose interaction. It would be beneficial to have support for a range of mobile platforms from portable notebook computers, to PDAs, to mobile phones.
Proposal: It seems unrealistic to make firm proposals in this area although the core requirement must be real-time support for interactive applications with the option of high-quality offline rendering.

10.2.2. Proposed Pipeline Contributions

Work on virtual humans will tend to be at the end of the pipeline and generally will be closely linked with 3D visualisation. It seems desirable to connect mini pipelines together to prove the effectiveness of the EPOCH infrastructure.

<u>Proposal</u>

On the assumption that presentations such as showcase 4, multilingual avatars, are of general interest, it will be beneficial to demonstrate that by using appropriate standards, new presentations can be created with only moderate work. The input will be 3D models of cultural heritage sites or objects along with metadata. The metadata will provide information about the model that can be used to generate behaviour for avatars that populate the scene and act as potential guides for visitors. Possible candidates for additional models would be Sagalassos from ETHZ and areas of medieval Norwich from UEA using the OpenSG framework with some natural language generation such as that provided by Brighton for showcase 4. (Could involve TU-BS, UEA, ETHZ, UNIGE, EFPL, Brighton.)

<u>Proposal</u>

Rapid development of virtual human characters, their clothing, and typical behaviour, will be needed if cost-effective presentations are to be developed. First steps for building a knowledge base of characters and clothing would enable different presentations to have access to high quality animations to suit particular requirements. The work would focus on the capabilities of VHD++, making accessible capabilities currently available only in bespoke animations. Methods for animation from motion-capture and motion synthesis would be exploited. (Could involve UNIGE, EFPL, UEA, Brighton, ETHZ.)

<u>Proposal</u>

In addition to the above proposals, where the main focus is rich use of virtual humans, it would be highly desirable to demonstrate end-to-end processing of cultural heritage resources. The models, metadata, and knowledge bases used above would be largely ad hoc, though confirming to standards as far as possible. In an end-to-end pipeline, the data at each stage would have been processed through the pipeline in a way that augments and preserves metadata in a realistic fashion, so that tools require little or no manual intervention to support integration of virtual humans at the final stage of producing a presentation.

10.2.3. Research Proposals

Avatar Standards for CH

Just as standards are required for representing 3D models of CH assets, to encourage reuse, so it is desirable to employ standards for avatar definition and animation parameters.

Some commercial formats are proprietary and inadequate for detailed animation of manual gestures. Formats, including MPEG-4 FBA are avatar-specific. Work is needed not only on use of existing formats and tools but also in the development of formats that handle CH metadata.

In order to achieve maximum reuse of resources, it will be highly desirable that motion files will be usable with a range of avatars. This will require the ability to retarget captured or keyframed motion, or to synthesise motion for a specific avatar definition.

The aim of the new tools would be in conversion between formats, preserving CH metadata, so that existing commercial and open source tools could be used to process avatars in the EPOCH pipeline.

Virtual Human Development Framework

Modern 3D VR systems relay heavily on the interplay of heterogeneous technologies. Because of that inherently interdisciplinary character the VR domain can be viewed as a melting point of various technologies which although complementary are non-trivial to put together.

Many of those technologies attract a lot of individual R&D interest but their integration is still not generally understood and there are barely any accepted guidelines and approaches for integration of those functional artefacts under the umbrella of a single consistent framework. In other words we now have many excellent exemplars of individual atomic technologies but still we lack a well understood and generally accepted strategy for putting them together so that they provide a whole which is bigger than the simple sum of its parts.

The missing element is an open source framework to glue them together. This would curb the complexity and make the resulting system machinery a consistent and seamless unity, leaving at the same time open handles and hooks for replacements and extensions.

The framework is to enhance the development of the interactive audio-visual real-time systems featuring real-time virtual character simulation for cultural heritage. A focus should be given on open source environment for complete and highly realistic virtual human simulation with body animation, skin deformation, facial animation, speech, and interaction in realistic 3D environments.

Avatar Development Tools

Avatars and their associated motion files are generally developed for a specific application. It is envisaged that EPOCH applications will have some common requirements for using avatars.

To develop compatible avatars, it may not be enough to provide methods for linking authoring tools, since it is often necessary to return to an earlier stage in the development pipeline, thus losing detailed work further on. New tools would either provide a complete open source system for developing EPOCH avatars, importing simple avatars from other tools, or would provide a system for merging changes made during different stages of development.

Populating 3D Environments

It is an expensive process to create and render 3D digital representations of CH objects and sites, whether representing existing artefacts or reconstructions. It is useful to add interest to such presentations using avatars, but this involves further expense and generally relies on bespoke tools.

EPOCH is developing standards for representing 3D models and for attaching rich metadata to support a range of presentations of models and enquiries about them. New tools would make it possible to render the 3D models in an environment where avatars, acting in response to the model metadata, could add interest and realism. Further, avatars could provide interactive guides that would respond to the visitor according to their language, interests, and past experience. An application of the tools would be automatic generation of populated environments as prototyped in Showcase 4.

Scripting Avatar Behaviour

The value added to EPOCH presentations by the use of avatars will depend on the avatars exhibiting realistic behaviour. It will be important that scripts can be generated in response to metadata about EPOCH sites and objects. This will require a sufficient repertoire of moves, gestures, and mouth movements for speech, along with avatars that are able to perform the repertoire.

Several EPOCH partners have developed XML-based notations for scripting dialogues and avatar movement. New tools would combine these notations into a common framework for use in the EPOCH infrastructure, along with at least one reference implementation. Such a notation will have wider application outside the CH area.

Mobile Avatar Platform

A benefit of using avatars as opposed to video for presentation of information is that animation data can be considerably more compact than video. The user has control of the presentation style and animation can be generated interactively. However, processing and memory resources on mobile platforms tend to be tuned for video and not always sufficient for full 3D animation.
In order to make avatars available on mobile platforms it would be desirable to track the development of devices and available multimedia tools. If available tools are adequate, then new tools would be developed to transfer avatar animation specified in EPOCH formats to such tools. Otherwise investigation would be made of the feasibility of porting existing avatar players to mobile devices.

10.3. References

de Heras Ciechomski, P. Ulicny, B., Cetre, R. And Thalmann, D. 2004. A case study of a virtual audience in a reconstruction of an ancient Roman Odeon in Aphrodisias. In Proc. VAST 2004 – Virtual Reality, Archeology and Cultural Heritage: 9-18.

Pablo de Heras Ciechomski, Sébastien Schertenleib, Jonathan Maïm and Daniel Thalmann, "Reviving the Roman Odeon of Aphrodisias: Dynamic Animation and Variety Control of Crowds in Virtual Heritage", VSMM '05, 2005

DigiCULT. 2004. Technology Watch Report 2. http://www.digicult.info.

Gaitatzes A., Christopoulos D. and Papaioannou G. 2004. The Ancient Olympic Games: being part of the Experience. In In Proc. VAST 2004 – Virtual Reality, Archeology and Cultural Heritage: 19-28.

Giannoulis, G. et al. 2001. Enhancing Museum Visitor Access Through Robotic Avatars Connected to the Web. In Proc. Museums and the Web 2001, Seattle, USA

Koutsonanos, D. Moustakas, K. Tzovaras, D. and Strintzis, M.G. 2004. Interactive Cloth Editing and Simulation in Virtual Reality Applications for Theater Professionals. In Proc. VAST 2004 – Virtual Reality, Archeology and Cultural Heritage: 37-46.

Kruijff, E., Conrad, S. Palamidese, P. Mazzoleni, P. Hasenbrink, F. Suttrop, M. and Kwon Y. 2004. Remote Virtual Guidance in Immersive Museum Applications. In Proceedings of VSMM2004, Gifu, Japan.

Leanizbarrutia, Igor; Behr, J.; Ortiz, Amalia; Linaza, María Teresa; Schnaider, M. 2003. New technologies-based diffusion of Cultural Heritage environments. 3rd IASTED International Conference on Visualization, Imaging, and Image Processing. September 8-10, 2003. Benalmádena, Spain

Magnenat-Thalmann, N. 2004. Modeling of Bodies and Clothes for Virtual Environments. Proc of CAVW 04.

Roussou, M. et al. 2001. Experiences from the Use of a Robotic Avatar in a Museum Setting. In Proc. VAST 2001 – Virtual Reality, Archeology and Cultural Heritage: 153-160.

Ryder G., Flack P. and Day A.M. 2004. Adaptive Crowd Behaviour to Aid Real-Time Rendering of a Cultural Heritage Environment. In Proc. VAST 2004 – Virtual Reality, Archeology and Cultural Heritage: 29-36.

Veltman K. 1998. Interfaces for Cultural Heritage.
http://www.sumscorp.com/articles/pdf/1998%20Interfaces%20for%20Cultural%20Heritage.pdf

Vlahakis V. et al. 2001. ARCHEOGUIDE: First results of an Augmented Reality, Mobile Computing System in Cultural Heritage Sites. In Proc. VAST 2001 – Virtual Reality, Archeology and Cultural Heritage: 131-139..

Wilkinson H., Jennings V.J., Glauert J.R.W. and Elliott R. 2004. A Speaking Avatar Toolkit for Cultural Heritage Applications. In Interdisciplinarity or the Best of Both Worlds. Selected papers from VAST 2004: 59-62.