Museums and
Public Archaeology Applications

The 5th International Symposium on Virtual Reality, Archaeology and Cultural Heritage VAST (2004) *Short Presentations*
K. Cain, Y. Chrysanthou, F. Niccolucci, N. Silberman (Editors)

Visiting Archaeological Sites with Our Mobile Phone: New Perspectives for Research in the Sixth Framework Programme. The Agamemnon case

D. Pian, A. Traverso, M. Villa

Darficlet, Univerity of Genoa, Italy

Recent advances in the availability of 3G – cellular phones with imaging and multimedia capability opens completely new horizons for the development and fruition of innovative IT tools in the domain of cultural heritage (archaeological sites and museums).

In the area of "cultural heritage", the Agamemnon Consortium has identified an opportunity to leverage on the most promising hardware and software technologies and use integrated approach to make the cultural heritage fully available and accessible so to transform the visit into a complete, culturally rich and long-lasting experience and to assist the site management in protecting heritage from a variety of risk factors (e.g. natural degradation, erosion and intentional damage).

Agamemnon has a number of distinctive features which represent clear advances over previous projects, e.g. **TOURBOT**, **ARCHEOGUIDE** and **PAST**.

TOURBOT's aim was to create a robot which could move in a museum being directed by the user over the Internet. It was restricted to museums, as an archaeological site would not offer the structured environment. In contrast "Agamemnon" is aimed at the user physically present at an archaeological site.

ARCHEOGUIDE aimed to support a user by supplying enhanced images; in that sense it had similar aims to "Agamemnon". "Archeoguide" relied on bulky Head-Mounted Displays to deliver the images and on tracking systems to localize the user. In contrast, "Agamemnon" will deploy a much lighter architecture and a much more sophisticated method of localizing the user.

The most innovative achievement of "Agamemnon" however will be the revolutionary way in which tracking will be achieved without the need for any further equipment. Instead, the position and orientation of the user will be determined in real-time through the analysis of the images shown in the user's handset. Furthermore, "Tourbot" and "Archeoguide" are predicated on a single end user, laboriously choosing his or her own path through the museum or archaeological site.

PAST went well beyond that in scheduling the path of a user depending on his or her preferences, time available etc. "Agamemnon" in turn, went beyond "Past" in recognizing that users are not alone in visiting a site and that therefore, it would be advantageous to plan the itinerary depending on the number, spatial distribution and even plans of other visitors. "Past" offered visitors a planned visit, but it did not enhance the visit with images of the past superimposed over the present (as "Agamemnon" does) nor it did take into account the presence of other visitors. "Past" depended on the technology of the day – portable PCs – which the museum or archaeological site had to have in store and to make available to each user. That increased the overall cost and therefore limited the exploitation potential.

In contrast, AGAMEMNON will exploit the functionalities of the user's own 3G handset. It will therefore do away with the need for costly and bulky extra equipment whilst providing functionalities well ahead of those provided with the earlier projects with their more costly options: schedules which take into account those of other users, position and orientation determined directly from the user's handset, capability of deployment even on the rough terrain of archaeological sites. Specifically, Agamemnon plans to exploit 3G mobile phones equipped with embedded cameras as the input/output device making use of graphical interfaces and voice-based commands to provide a visitor to a site of historical interest with personalised information , while at the same time contributing to the preservation of cultural heritage.

There are some activities the visitors can perform from home before visiting the site, in order to prepare for the visit and let Agamemnon be more effective during it. Such activities are related to supply a set of personal information required in order to let the system build a specific profile of the visitor (we call this "static profile"). So Agamemnon will offer a web site where people can log in and provide their preferences for preparing their next visit.

How will this process take place? At the beginning, users will have to choose the language (Agamemnon will be available in English, Italian and Greek, at the beginning), a login name and a password they will use to authenticate themselves, when they will be on-site, through their mobile phones. After that, a main page will provide a list of options on various areas. Then, the user will be directly asked what kind of interests he has in different cultural aspects linked with the visit.

The visitor's interest scores will be used by the Visitor Optimiser to determine which monuments should be included in the visit. Each monument will be given a score based on how relevant it is and the Visit Optimiser will attempt to choose a route that is as interesting as possible based on these scores. The visitor's interest scores will also be used to determine which Additional Topics should be displayed when the visitor is looking at a particular monument. Finally, after supplying all the information, it will be possible to download the mobile application directly into the phone. The application will be already configured for that specific phone model. In case a user doesn't have the time or the possibility to connect to the web-site before starting the visit, as soon as he start using Agamemnon in the site, the system will prompt directly such questions.

During the site visit the system will be displayed directly into the small screens of the 3G cellular phones, as well as all the interaction with the user will be through his mobile.

During the visit Agamemnon will be performing various options suggesting a personalised visit path and providing personalised multimedia material on the visit.

Anyway, Agamemnon should not force a visitor to follow its route, but rather provide "suggestions" and be ready whenever visitors demand for additional information.

With the main homepage the system welcomes the visitor, who can decide to start with the suggested visit, to go directly to a monument he prefers or to take a picture to a monument he likes, and Agamemnon will recognise it and propose to visit it.

The monument overview provides a quick description of the next monument Agamemnon suggests to visit or the one photographed by the visitor. A picture is shown in order to help the visitor to identify it, and also the system provides some easy indications on how to get there.

A user can decide to visit a specific monument in any moment during his visit. There are two possibilities, either choose the monument from a scroll list including preview pictures, or take a picture to it and let Agamemnon identify it.

If Agamemnon is not able to identify a monument directly, the system will ask the visitor to identify the monument by recognising it from a scroll-down list with preview pictures.

When a visitor decides to visit a monument, the initial screen shows a list of pre-selected topics and a picture of the monument. The system will provide multimedia information on a specific topic of the monument (art, culture, history, etc..).

A topic is a set of multimedia information (descriptions, pictures, movies, speech) on a specific topic on a monument. It should be reminded that the information displayed depends on a specific visitor's profile, so the same basic information could be presented in different ways to different users.

When the user clicks on "Conclude", Agamemnon will terminate the visit and close the connection.

Agamemnon aims at contributing to Preservation of cultural heritage thanks to the **Preservation Monitoring Module:** it will use images collected through interaction with the visitors to offer site managers tools to monitor the status of conservation of the site (e.g. monitor erosion and deterioration of artefacts; detect damages; etc.), allowing an automated and constant monitoring activity on the whole archaeological area without requiring the intervention of specialised and dedicated personnel, and producing a more timely reaction in case of significant changes in the monument's preservation state.

In un-guarded sites Agamemnon could be the only mean of preserving monuments, providing a cheap and convenient way of remote monitoring of the site: in fact it is just sufficient to have someone at the organisation responsible for the site receiving the alarms from Agamemnon. This situation is also reflected by the actual condition of the two pilot sites (Paestum, Mycene) where a security system is not present (except on-site guardians), and where past attempts to install fixed cameras failed due to various reasons.

The images taken by the visitors are collected in chronological order, and this will be useful to analyse the state of the site in time.

Thanks to the Imagines System, Agamemnon will be able to understand which monument each picture is taken from. In this way, it will be possible to perform an auto-categorisation of the security images, and site personnel will be enabled to have an instant snapshot of the situation of each monument in the site.

The system will have a "warnings" section, and site personnel will be enabled to mark an image as "warning"; as consequence they may decide to perform an in-site analysis to check the real current state of the monument.

The 5th International Symposium on Virtual Reality, Archaeology and Cultural Heritage VAST (2004) *Short Presentations*
K. Cain, Y. Chrysanthou, F. Niccolucci, N. Silberman (Editors)

On the Road with the Iberians

A. Herranz[1], D. Pletinckx[2], H. Tency[2], A. C. Ruiz Rodriguez[1], M. Molinos[1]

[1]Centro Andaluz de Arqueologia Ibérica, Paraje Las Lagunillas, 23071 Jaén, Spain
[2]Ename Center for Public Archaeology and Heritage Presentation, Abdijstraat 13-15, B-9700 Oudenaarde, Belgium

Abstract

This paper presents the results of work accomplished a bursary within the context of EPOCH, the European Network of Excellence on Intelligent Cultural Heritage. This bursary was linked to the theme of Cultural Routes, and creates a first system within the "Viaje al Tiempo de los Iberos" route. The system interprets an Iberian burial chamber, and gives virtual access to the archaeological site.

Categories and Subject Descriptors (according to ACM CCS): H.3.3 [Information Search and Retrieval]: Retrieval models, H.5.2 [User Interfaces]: Graphical user interfaces (GUI), H.5.4 [Hypertext/Hypermedia]: Navigation

Introduction

Cultural Routes, especially on archaeological heritage, have specific needs. Archaeological sites can be in the countryside, far away from the cities or villages, and are prone to damage by unsupervised visitors or vandals. In addition, the use of conspicuous informational panels can jeopardise the authenticity of the site. This paper describes a prototype system that solves both problems by creating virtual access to an extremely well preserved burial chamber of the Iberical culture.

About the Iberian Route

The archaeological site of Cámara de Toya is part of the "Viaje al Tiempo de los Iberos" (Route to the Time of the Iberians), a project developed by Deputation of Jaén [Via] in close collaboration with the "Centro Andaluz de Arqueología Ibérica" [CAAI].

The main objective of this initiative is to present the rich Iberian culture, archaeological remains of which the province of Jaén possesses to an exceptional degree. This project attempts to join archaeology, heritage, and history in the context of a cultural route.

The "Viaje al Tiempo de los Iberos" (see Figure 1) is structured as a journey from the time of arrival of the Romans on the Iberian Peninsula, during the second Punic war (203 BC.)—which marked the submission of the Iberian people of the upper Guadalquivir, going back into time to the splendour of the Iberian aristocratic society in which an Iberian prince controlled a large territory (IV BC.), to arrive finally at the origin of the Iberian culture, when the first towns or Oppida developed (VII-VI BC).

This cultural route seeks to present a full picture of who the Iberians were, where and how they lived, and their basic beliefs. The route also highlights the different types of archaeological sites that exist in the Iberian culture and the material culture that is left by the Iberians. The route has seven stops and can be followed from beginning to end in two and a half days.

Figure 1: Viaje al Tiempo de los Iberos.

The Deputation of Jaén supports the signposting and promotion of this route, maintains the infrastructure, ensures accessibility of the sites, and organises exhibitions and supports excavations and studies related to the route.

The burial chamber of Toya

One of the stops on this route is the burial chamber of Toya (see Figure 2), discovered at the beginning of the twentieth century and believed to be the tomb of an aristocrat from the oppidum of Tugia, now called Toya [Rui99].

The chamber has a rectangular floor plan and five compartments in three longitudinal aisles, and contained many objects. A life-size copy of the chamber has been made in the Archaeological Museum of Jaén, containing replicas of the original grave finds.

The archaeological site is situated in the middle of agricultural fields. It has been fenced and an access ramp has been added for visitors. Close to the site, an interpretation centre and parking lot have been built.

Figure 2: The burial chamber of Toya.

Figure 3: Interpretation Centre.

Figure 4: The user interface of the Toya system.

Implementation

The Camera de Toya system is based on TimeScope 3 software [PSC03] and consists of a virtual walk around and in the burial chamber. From this virtual walk, links are provided to relevant historical and archaeological information, such as the context of the Iberian culture, background information on the burial chamber, and its on discovery, excavation, and restoration. The language of the current prototype is Spanish but multiple languages will be available in the final version.

The rationale of the system is that the visitor builds his own story by exploring the provided information based upon own interest. In this way, the system adapts to the level of knowledge and interest of the visitor without complex user interfaces or multiple versions of the presentation.

Another main feature of the system is the use of a visual language, based upon panoramic virtual tours, so that the visitor can explore in a simple and natural way large quantities of information without any menu structure.

The system is intended to be used in the Interpretation Centre to give background information to the visitor before visiting the chamber. In this way, a physical visit to the site becomes a much richer experience as the virtual tour links clearly to the physical visit.

Secondly, the system can be used whenever a physical visit to the site is impossible. This can be the case for large groups, or when no personnel are available to open the site which is at a few hundred meters from the Interpretation Centre. Unattended visits pose the danger of vandalism to the site, as has happened in practice.

Thirdly, the system can be used in combination with the replica at the Jaén museum, creating a link with the archaeological site, and encouraging visitation there.

It is expected that the Camera de Toya system will eventually become part of a computer aided cultural route that supports and links the multiple sites of the "Viaje al Tiempo de los Iberos". This computer aided cultural route consists of a personalised Internet portal, local presentation systems and a central user database [ODP04]. The portal serves as starting point where the sites, their facilities and the routes in between are listed, so that the visitor can plan his journey throughout the cultural route. Listing local facilities helps to create economic impact in the region of the cultural route, especially in rural areas [BW04]. After visiting a site, the content of the site system becomes available at the personalised portal, and forms a kind of logbook of the visitor on his journey **[Ple02].**

Figure 5: The input interface of the Toya system.

Input and updating of the information does not require technical skills (see Figure 5) and is based on a clear visual language, so that archaeologists and historians understand the content can create and maintain such systems.

Acknowledgements

Many thanks to Heidi Tency, Truus Helsen, Lars De Jaegher and Daniel Pletinckx of the Ename Center who were of great support during the realisation of this project. This project is based upon the TimeScope3 software that was commissioned by the Provincial Museum in Ename (directed by Marie-Claire Van der Donckt), with partial funding from the Flemish Ministry of Culture.

References

[BW04] BRIEDENHANN, J., WICKENS, E.: Tourism Routes as a tool for the economic development of rural areas –vibrant hope or impossible dream? Tourism Management 25 (2004), 71-79.

[CAAI] The Andalusian Center for Archaeology http://www.ujaen.es/centros/caai

[ODP04] R. OWEN, D. BUHALIS, and D. PLETINCKX : Developing the tourism aspects of a cultural route, *VAST2004 Proc.*

[Ple02] PLETINCKX, D.: Computer Aided Cultural Routes, Proc. of 3D Virtual Heritage workshop, Geneva, (2002)

[PSC03] PLETINCKX D., SILBERMAN N. and CALLEBAUT D.: Heritage presentation through interactive storytelling: a new multimedia database approach, *J. Visual. Comput. Animat. 2003; 14: 1–7*

[Rui97] RUIZ RODRIGUEZ A., The Iron Age Iberian Peoples of the Upper Guadalquivir Valley. *The Archaeology of Iberia,* Routledge, New York, 1997; pp. 175-191.

[Via] Viaje al Tiempo de los Iberos: http://www.promojaen.es/

The 5th International Symposium on Virtual Reality, Archaeology and Cultural Heritage VAST (2004)
K. Cain,Y. Chrysanthou, F. Niccolucci, N. Silberman (Editors)

On-a-Slant: A Mandan Village in Virtual Time and Space

J. T. Clark, D. G. Snider, A. Bergstrom, R. Frovarp, D. Reetz, B. M. Slator

Archaeology Technologies Laboratory, North Dakota State University, Fargo, ND, USA

Abstract

This paper summarizes a project designed to create a virtual site simulation for public entertainment and education. The use of high detail, accurate coloring, and dynamic effects provide a high level of realism. The use of stereo 3D projection further heightens the realism by creating an immersive experience.

1. Introduction

The United States recently began a bicentennial celebration of the epic journey of exploration undertaken by Merewether Lewis and William Clark, from 1804 - 1806. Lewis and Clark led the „Corps of Discovery" from St Louis, up the Missouri River, across the Rocky Mountains, westward to the Pacific Ocean and back again. Each state through which the Lewis and Clark party traveled will be hosting a major "Signature Event" to commemorate the bicentennial of the expedition. In North Dakota, the event took place October 22-31, 2004. As part of that event, the authors and a group of dedicated and talented colleagues created a virtual village that was occupied by Mandan Indians about two decades before the Lewis and Clark journey. This paper discusses the creation of that virtual village, which will be shown at the VAST conference.

2. The site

By the late 18th century, the Mandan Indians lived in seven fortified villages along the Missouri River in central North Dakota. One of those Mandan villages was On-A-Slant, which was established in the latter half of the 16^{th} Century. In 1781, a devastating smallpox epidemic struck, that, coupled with increased raids by the nomadic Sioux tribe, forced the survivors of On-a-Slant and the other villages to move farther upriver.

Twenty years later, on August 20, 1804, William Clark would write of the ruins of On-a-Slant in his notes: „Saw an old Village of the Mandans below the [Heart River]. [A]ppear to have been fortified." He was told that the people who had lived there moved farther upriver [Jen03: 59]. Six days and 60 miles later, as the Corp of Discovery approached the mouth of the Knife River, they encountered the two remaining Mandan villages near four villages of friendly Hidatsas. The Corps wintered nearby, and it was the friendship of the Mandan and Hidatsa Indians that helped ensure the Corps' survival through the harsh winter. On their return voyage 22 months later, the Lewis and Clark expedition again came to the confluence of the Heart and Missouri rivers. William Clark used the remains of On-a-Slant Village to elicit more information about the Mandan people from one of their chiefs.

Just as Clark used On-a-Slant to learn more about the Mandan, we have used the village as an entry point to explore the traditional life of the Mandan people. Today, the On-a-Slant Village site is an interpretive park operated by the State of North Dakota. The park includes replicas of six earth lodges and a museum. The lodge replicas at On-A-Slant make it one of the few places where visitors can get a feel for what an earth-lodge village might have been like. That experience is highly limited, however, in visual scope, context, and information. Furthermore, the practice of building physical 'reconstructions' at actual sites is controversial. By employing state-of-the art technology together with scholarly research and artistry, we have created a virtual model of what the entire village may have looked like two centuries ago. Certainly we can never know 'exactly' what the past was like, but we *can* produce models accurate to the available evidence. Such virtual models can provide heritage monuments and museums with interpretive visualizations that are cost-effective and nondestructive alternatives to physical replicas.

3. Virtual site construction

The Fort Abraham Lincoln Foundation, which provides funding for the On-a-Slant site, contracted the Archaeology Technologies Laboratory to create the virtual On-a-Slant Village exhibit for the North Dakota Signature Event. The interpretive model produced is based on intensive research, logical extrapolation, and analogy. We used a variety of old documents (maps, traders journals, ethnohistorical descriptions, etc.), cultural and historical studies, excavation reports from Slant and other villages appropriate to the general time and place, paintings and drawings of slightly later

Mandan villages and people (e.g., by Bodmer, Catlin), physical reconstructions of earth lodges made by others based on prior research, and a vegetation reconstruction.

Our previous work on interactive virtual worlds provided valuable guides for the On-a-Slant project. Even so, the project covered much new ground and encountered many new pitfalls. Despite those many problems, the final product is a virtual world of very high realism.

The three-dimensional (3D) computer-based model is a selected slice in time, or „simulation slice." For this simulation slice we show the village at the height of its prosperity, in 1776, when it was home to approximately 1,000 people living in up to 86 earth lodges (the distinctive house form in the region). We modeled the village in late summer to show the gardens and highlight the agricultural subsistence of the Mandan.

The virtual world was created using Maya Unlimited by Alias, Inc. After many months of background research and 3D modeling, we developed a preliminary virtual village, covering 33,500 sq. m, set in a landscape of 559,500 sq m. We then established a camera path for the fly-through, a narration script was written, and the two were coordinated. The camera rig consisted of two parallel cameras offset slightly to create eye separation. These were controlled by a transform node attached to the camera path. After the final modeling touches were completed, the world was rendered to video with one of the cameras. A second rendering was then made with the second camera. These two renders, one for each eye, yield the final stereographic video when played on two synchronized DVD players. Polarized glasses give the immersive 3D effect. The complete video is a little under 10-minutes long. It can also be viewed in monographic format, as shown at VAST 2004, or on the internet at this link: http://onaslant.ndsu.edu/videoclips/.

The video begins with a scene looking up and across the river. The view (camera) moves upstream, past the village, before circling around the gardens. The video path continues around the defensive palisade and into the village. It moves through the village, around the densely packed earth lodges, through the central plaza, and into one of the lodges. After moving about the lodge interior, the camera exits through the smoke hole at the top, then starts moving upriver, in the direction taken by Lewis and Clark, as the video ends. Narration and appropriate ambient sounds are heard throughout in surround sound.

We used laser scanners to create 3D models of a small set of artifacts recovered from the site, and placed those models within a few selected scenes. When viewed in stereo, these artifact surrogates appear to come out of the screen and 'float' in front of the viewer, rotate, then diminish again to proper proportion. For example, we see a woman scraping a deer hide and hear the scraping sound, captured from a real hide scraping. The scraper, made from flaked chert in an elk-antler handle, leaves her hand and enlarges as it moves toward the camera. The scraper separates from the handle and rotates, seemingly in front of the screen. The stone scraper returns and to the handle and the tool moves back into her hand.

Human figures are sparsely scattered through the world. The premise for having such a small population at the village—which is done to improve visibility of the village and for ease of modeling—is that most of the residents are still away on a late summer buffalo hunt but will soon return. The people are created in full 3D. Only a few show movements, as do a few animals, animated through the use of a skeletal rig. In the future, we hope to animate human movements through the application of motion capture data to enhance realism.

A key to realism and sense of place is the application of dynamic effects. These effects were created with Maya Dynamics, which uses mathematical formulas to calculate real-world physical actions. For example, a revised ocean shader was used to simulate the flow of the river, and fluid effects were used to create the movement of a boat on the river. More use of dynamics was originally planned but were later removed in favor of shorter render times. The use of dynamics can greatly enhance a scene, but also significantly increases render time, which was a constant impediment to the project.

Shadows, especially those from ray traced lighting, also significantly increase render time, so they were used only when absolutely necessary (e.g., lighting the river, which required ray tracing to calculate reflections and refractions). For the rest of the scenes, directional lights utilizing depth-mapped shadows, point lights, and low intensity ambient light were used.

Rendering was done with Maya's built in Software Renderer. Over 30,000 frames were needed to produce the stereo video at 30 frames/second, with two runs for stereoscopic viewing. The render time ranged from 6 to over 120 minutes per frame depending on the dynamics and animation used in the scene and the number and type of objects in the field of view. Only by using a large number of dedicated computers (at maximum 182 processors were employed) in multiple clusters could we accomplish the grand feat of rendering so many frames in a relatively short period (ca. three weeks).

4. Conclusion

With the conclusion of the Circle of Cultures Signature Event, the On-A-Slant virtual exhibit was reinstalled as a permanent display at the Fort Abraham Lincoln State Park Museum. Our goal in this project was to make a place and time in history come alive for viewers, if only for a several minutes. In so doing, we can satisfy our obligation to bring cultural heritage to the general public in a way that is engaging as well as educational.

Acknowledgements

This material is based upon work supported in part by the National Science Foundation under Grant No. 0224671, and by funding from the Fort Abraham Lincoln Foundation. Alias, Inc. provided the 3D authoring software.

References

[Jen03] Jenkinson, C. *A Vast and Open Plain*. State Historical Society of North Dakota, 2003.

The 5th International Symposium on Virtual Reality, Archaeology and Cultural Heritage VAST (2004) *Short Presentations*
K. Cain,Y. Chrysanthou, F. Niccolucci, N. Silberman (Editors)

A Speaking Avatar Toolkit for Cultural Heritage Applications

N. Wilkinson, V. J. Jennings, J. R. W. Glauert, R. Elliott

School of Computing Sciences, UEA, Norwich, UK

Abstract

This paper describes development of a toolkit for speaking avatars to be embedded in virtual models of cultural heritage sites. The avatars must make efficient use of system resources, be simple to develop and appear highly realistic. The toolkit will generate accurate mouth movements for a speaking and gesturing avatar that will act as a virtual guide for visitors to virtual cultural heritage sites, adding interest and effectiveness to such models.

Categories and Subject Descriptors (according to ACM CCS): H.5.1 [Information Interfaces and Presentation: Multimedia, Information Systems]: I.3.3 [Computer Graphics: Picture/Image Generation]

1. Introduction

The EPOCH showcase project at the University of East Anglia demonstrates a set of tools that allow heritage sites to be recreated in a virtual environment quickly and easily. The work builds on tools developed within the EU Charismatic project (http://www.charismatic-project.com) [FWB*01]. The sites will be populated with virtual humans, or avatars, to give an enjoyable learning experience. Visitors can explore the virtual environment interactively with the assistance of a talking avatar that can function as a guide. The toolkit aims to add realistic speech to the scriptable gesturing avatars developed within the EU ViSiCAST project (http://www.visicast.cmp.uea.ac.uk) [EGJK04] for animating natural sign language used by deaf people.

1.1. Realistic speaking avatars

Speaking avatars, or *talking heads*, have been developed for a number of applications with different design objectives. The Synface project aims to aid hearing impaired people understand telephone conversations by using speech recognition to drive a synthetic face [KFS03]. Other applications, such as support for speech therapy, require the talking head to be highly realistic. Video realistic speech synthesis [TCG*03] involves concatenating visual images of a real person. Work is in progress to integrate video synthesis with 3d avatars, but current systems are currently computationally expensive.

1.2. Specification of the speaking avatar

Avatar guides should fulfill the following specification:

- They should be multi-lingual. In due course, ViSICASTbased sign language support can be provided.
- They should be quick and easy to develop since they will ultimately be made by non-technical people.
- They should both sound and look as realistic as possible.
- They should consume minimal system resources as the 3d worlds which they inhabit are computationally intensive. Even stricter requirements will apply for deployment on personal digital assistants (PDAs).

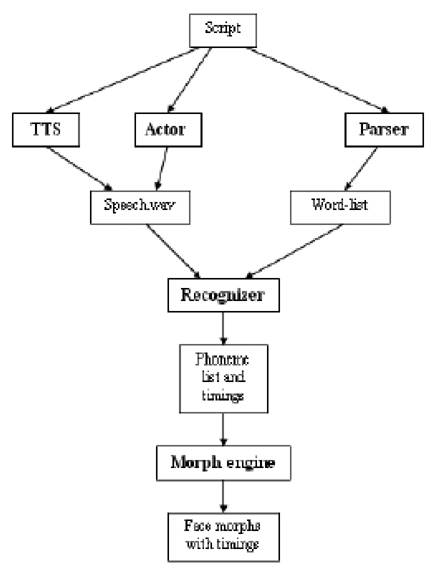

Figure 1: Processing stages for a talking avatar.

2. Speaking avatar architecture

Figure 1 shows the proposed architecture which supports multiple avatars each with a personalized voice. Allowing the speech to be generated by a Text-To-Speech (TTS) synthesizer enables the avatars to be interactive. A script for the avatar is passed to an actor or a TTS synthesiser and a parser. The actor or TTS synthesizer generates a wave file containing speech for the avatar. The parser removes punctuation, adds pauses where necessary and produces a word list. The speech recogniser aligns the speech with the word list and produces a list of phone states and their timings.

Using a recognizer allows the phone set to be independent of the any TTS system and provides timing information at the phone-state level. This is useful since articulatory gestures such as stop consonants can occur within phones. The innovation in the toolkit is in the use of phone-states to achieve better handling of co-articulation than is possible using complete syllables.

3. The morph engine

The morph engine takes the list of phone-states with corresponding timings and smoothly interpolate between them before mapping them to a set of morphs applied to the avatar's face for each frame. Each phone-state has an associated set of Phone-State-Targets (PSTs), one for each articulator modelled, specifying the mean position of the articulator and the variance about the mean, estimated using the Mocha database [WH00].

Trajectories for each articulator are computed using the PSTs. The closeness of fit of an articulator trajectory to the PST mean depends upon the variance of the PST and the value of the surrounding PSTs. A low variance for a PST constrains the trajectory to come close to the PST mean. Allowing the surrounding articulators to influence the trajectory allows co-articulation between phones to be modelled.

3.0.1. The morph set

Figure 2 shows the reference face, a, and the six morphs, bg that can currently be applied to it. Initially 6 morphs are used, b: jaw-open, c: lips-rounded, d: lips-pursed, e: lipscompressed, f: tongue-tip-up, g: tongue-tip-forward. The primitive face shapes are drawn from the Facial Action Coding System (FACS) [EFH02].

Once the trajectories of the articulators have been computed they need to be mapped into the morphs. The avatar face shapes are generated by combining of various degrees of the six basic morphs:

- jaw-open = upper incisor Y - lower incisor Y
- lips-rounded = upper lip Y - lower lip Y
- lips-pursed = (upper lip X + lower lip X) / 2
- lips-compressed = -upper lip Y + lower lip Y
- tongue-tip-up = tongue tip Y
- tongue-tip-forward = tongue tip X

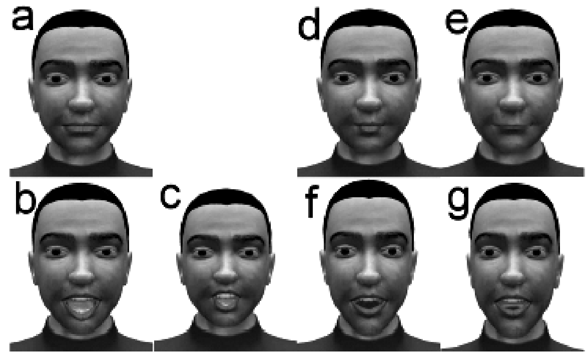

Figure 2: The six morphs used for animation.

3.1. Conclusions

Work to date has fulfilled much of the specification. By limiting the number of morphs used to a small number, new avatars will be quick and easy to develop. Early work on the visual speech synthesis is promising. The timing of speech gestures seems highly accurate and it does not appear that speech has simply been dubbed. The small database of 135 phone targets and simple co-articulation model allows the system to be very resource efficient both in memory and processing time.

References

[EFH02] Ekman P., Friesen W., Hager J.: *The Facial Action Coding System*. Weidenfeld & Nicolson, 2002. 2

[EGJK04] Elliott R., Glauert J., Jennings V., Kennaway J.: An overview of the SiGML notation and SiGMLSigning software system. In *Fourth International Conference on Language Resources and Evaluation, LREC 2004* (Lisbon, Portugal, 2004). 1

[FWB_01] Flack P., Willmott J., Browne S., Day A., Arnold D.: Scene assembly for large scale reconstructions. *Proceedings of VAST* (2001). 1

[KFS03] Karlsson I., Faulkner A., Salvi G.: SYNFACE - a talking face telephone. In *Proc EuroSpeech 2003* (2003). 1

[TCG_03] Theobald B., Cawley G., Glauert J., Bangham J., Matthews I.: 2.5d visual speech synthesis using appearance models. In *British Machine Vision Conference* (Norwich, UK, 2003), vol. 1, pp. 43–52. 1

[WH00] Wrench A., Hardcastle W. J.: A multichannel articulatory speech database and its application for automatic speech recognition. In *5th Seminar on Speech Production: Models and Data* (2000).