IST-2002- 507382

EPOCH

Excellence in Processing Open Cultural Heritage

Network of Excellence

Information Society Technologies

D.2.4.6 (Final): Showcase 6 "Avatar based Interactive Storytelling"

Due date of deliverable: 29 April 2005
Actual submission date: 28 April 2005

Start date of project: 15/03/2004
Duration: 4 Years

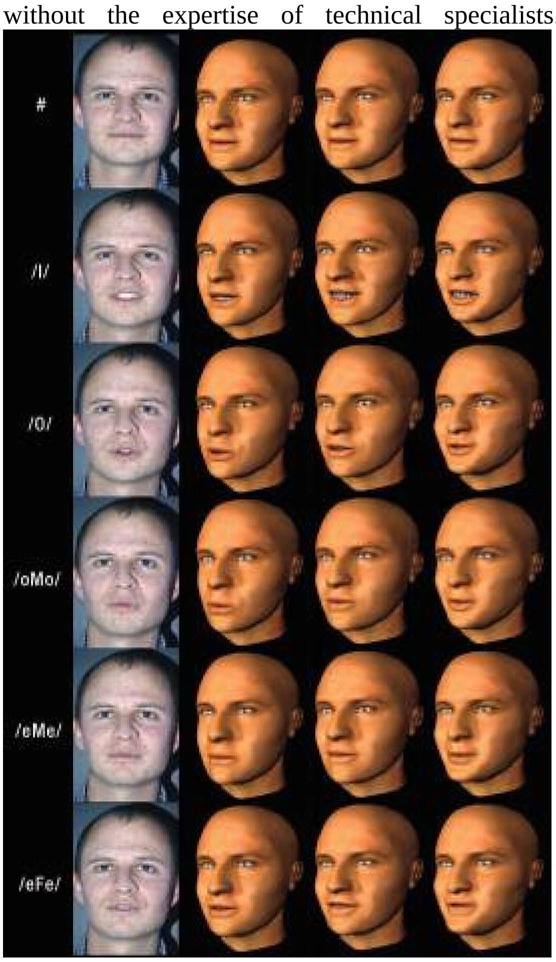
Ename Center

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)	

EPOCH Showcase 2.4.6
AVATAR-BASED INTERACTIVE STORYTELLING

The Ename Center for Public Archaeology and Heritage Presentation, Belgium
ETH Zürich, Computer Vision Laboratory, Switzerland
MIRALab, University of Geneva, Switzerland
with the support of Prof. Em. Dr. Arch A. L. J. Vande Walle

Storytelling is a familiar and effective way to convey information to a general audience, especially in the field of Cultural Heritage.
Yet with the increasing use of standardized interactive applications in museums and historic sites, it is difficult to appeal to a wide range of visitor interests.
This problem can be solved through the use of interactive storytelling, which allows the visitor to choose from a large selection of subjects and themes to create a personalized "story," appropriate to his or her interests and the time available for the visit.
Presentation systems featuring interactive storytelling should be easy to update and maintain, without the expertise of technical specialists.



Text-based interactive story telling applications, driven by a relatively simple XML or database application, allow museum and site staff to update the content by merely typing in texts. Recorded narration is generally more effective than text in conveying information to the visitor, but its use in interactive storytelling has a major drawback concerning updating : changing minor parts of the content requires major investments in time and money (studio recording sessions, editing, digitizing) which often results in discouraging frequent updates and jeopardizing the entire update process.

By adding synthetic voice capabilities to the interactive storytelling system, this problem of updating time and expense can be effectively addressed. Text-to-voice translation software makes it possible to create narration from the keyboard. Simple software tools can be used to alter intonation or stress significant words, making the synthetic speech more lifelike. The sound files can be easily exchanged in the application and they can provide narration in several languages.

A Virtual Guide

An additional element can greatly enhance the effectiveness of an interactive storytelling system: the use of an avatar to serve as personal guide and storyteller for each visitor.
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. And this element too can be integrated into a system that is easily updated.

The Archaeological site of the Abbey of Ename in Belgium was used as a test site for a showcase that demonstrates Avatar- Based Interactive Storytelling.

In the prototype system, the Avatar's speech and animation is fully synchronised with the synthetic voice, derived semi-automatically from written text in storytelling database. This integration can provide easily updatable multi-lingual, interactive explanations of major points of archaeological, cultural, and scientific interest in the form of self-authored "stories" conveyed by a virtual guide.

The Ename showcase also utilizes Virtual Humans to "populate" digital reconstructions of the site and to illustrate questions about archaeology, the abbey and the monks who lived there. Issues of alternative architectural reconstructions techniques also form a research priority for this showcase.

Technical specifications

The showcase application is based upon the TimeScope 3 software, developed by the Ename Center and IBM, which is an Open Source XML and Java environment, build upon an Apache Tomcat server. Data entry for content specialists uses standardised forms to input texts and link images, animations and sounds. The text-to-speech software called RealSpeak Solo is developed by ScanSoft Inc.

The synchronised avatars are a development of the Computer Vision Lab of ETH Zürich, and use standardised emotions.

The virtual humans are created and animated by MIRALab.

EPOCH showcase 2.4.6
AVATAR-BASED INTERACTIVE STORYTELLING
Creating Animated Characters

MIRALab

I. INTRODUCTION

The goal of the EPOCH applications is to bridge the gap between the IT sector and the target audience of cultural heritage professionals working in museums, monuments and sites, by demonstrating certain results of integration research as working prototype applications with real data. Therefore, several showcases have been created to stimulate the imagination and provide experience of current results.

In this text, we describe the technical aspect of the showcase 6, "Avatar based interactive storytelling". The archeological site of Ename and its abbey has been chosen for the 3D simulation of the daily life inside the Abbey during the 12th century.

The main objectives of the showcase 6 are focused in three areas:

a. Reconstructing the physical structures and spatial environment. (Ename Center)
b. Utilizing virtual humans to illustrate social structure, trades, and daily activities. (Miralab, University of Geneva)
c. Utilizing a digital Avatar as a "virtual guide". (Eidgenossiche Technische Hochschule, Zurich)

The intention of this project is to integrate these components in a "virtual" walk-through composed of a series of several interlinked panoramas at various points in the archeological site of Ename and its abbey.

In the following sections, we provide an overview of virtual human technologies being employed and we present early results based on the ongoing research. This description is focused on the research methodology we used to develop the complete chain of processes for the making of bodies and clothes. The work is divided in three steps: 1) virtual humans' creation, 2) virtual humans' animation, 3) virtual cloth simulation.

II. RESEARCH METHODOLOGY

1. 3D body creation

With virtual bodies, we consider mainly two issues. The first issue is that bodies should correspond to real human body shapes. The second issue is that virtual bodies should have an appropriate 3D representation for the purpose of cloth simulation and animation.

The first phase of our methodology is based entirely on the generation of human body models that are immediately animatable by the modification of an existing reference generic model. Our modeler builds on a template body model that has been designed primarily for real-time animation applications. The model encapsulates all required structural components—bones and skin attachment data that allow to animate the model at any time during the synthesis, through skin attachment recalculation; landmarks and contours through which it remains measurable.

Our approach is based on examples [1] and consists of three major parts. First, each example from the 3D range scanner is preprocessed so that the topology of all examples is identical. Second, the system that we call the "modeling synthesizer" learns from these examples the correlation between the parameters and the body geometry. After this learning process the synthesizer is devoted to the generation of appropriate shape and proportion of the body geometry through interpolation.

a. Data acquisition and preprocessing

• 3D scanned data

Example models were obtained from scanned bodies (50 for each) of European adults, acquired mostly from the Tecmath scanner [2]. Texture data was not available and is not within the scope of this work. Each of the 3D scanned data was measured at various desired parts and was annotated with its measured values.

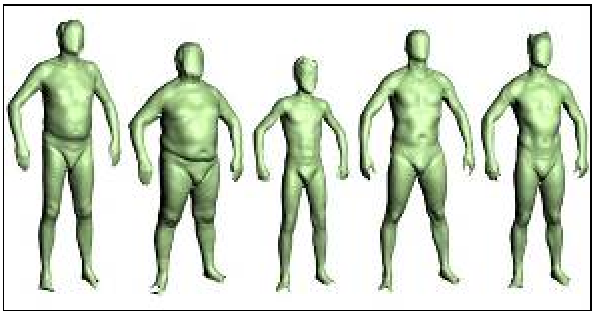


Fig. 1. The Scanned models

• Reference model

As shown in Fig.2, we use a two-layer model for the reference—skeleton and skin mesh without intermediate layer representing the fat tissue and/or muscle. The skeleton hierarchy that we use is composed of 33 joints including the root [3]. The template mesh is essentially a set of horizontal contours and vertical lines, forming a set of Bezier quad-patches. Such a grid structure enables us to take all supported measurements— girths and lengths—immediately, at any time during the process. Now that we have the skin and the skeleton, we must perform a process called skinning or skin attachment so that a smooth skin deformation can be obtained whilst the joints are transformed. Most character animation in interactive applications is based on this geometric skeletal deformation technique[i], which is also employed here.

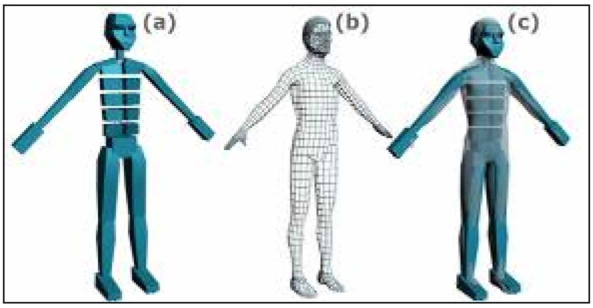


Fig. 2. Skeleton hierarchy; (a) The template model; (b) Skinning setup; (c)

• **Preprocessing**

The body geometry is obtained by conforming the template model onto each scanned models. An assumption made here is that any body geometry can be obtained by deforming the template model. A number of existing methods such as [4] and [5] could be successfully used. In this work, we adopt a featurebased method presented in [5]. The basic idea is to use a set of pre-selected landmark or feature points to measure the fitting accuracy and guide the conformation through optimization. There are two main phases of the algorithm: the skeleton fitting and the fine refinement. The skeleton fitting phase finds the linear approximation (posture and proportion) of the scanned model by conforming the template model to the scanned data through skeleton-driven deformation. Based on the feature points, the most likely joint parameters are found that minimize the distance of corresponding feature locations. The fine refinement phase then iteratively improves the fitting accuracy by minimizing the shape difference between the template and the scan model. The found shape difference is saved into the displacement map of the target scan model.

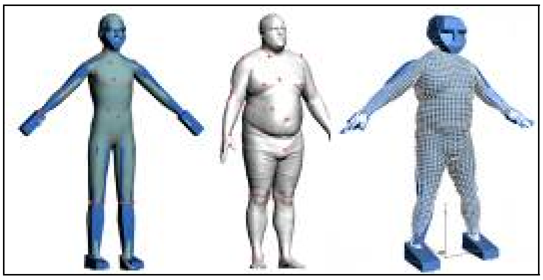


Fig.3 Skeleon fitting and fine skin refinement

b. *Head modelling*

For the 3D head modelling we used the face-building techniques [12], which allows you to map a 3D face and shape a head of your self or anyone using two photographs (Front and side).

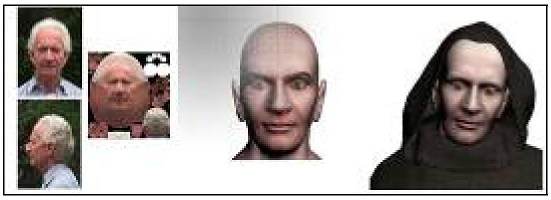


Fig.4 From 2D pictures to 3D model

To complete the modelling of a head it is essential to include hair. The colour and style of hair is often an indicator of gender, race and individuality. The hairdo is the same for all the monks, but not for the novice. They have a 'tonsure'; only a ring of short hair is left. The hair model we created includes these features and we simulated the 3D hair using our homemade software [13].

Fig.5 From 2D pictures to 3D model

2. Body Animation

According to our ‘scenario’ the 3D characters have to give the impression that they become ‘alive’ and they emerge from the past to tell their stories. Although realistic animations of short duration and quality were entirely carried out by keyframing, the motion capture remains the technique privileged to obtain realism. We used Optical Motion Capture based on markers [6, 7]. Ideally, the more the proportions between the virtual character and the subject agree, the better is the reconstitution. Well conceived, a motion capture recording can save a lot oftime and give a very realistic result.

Fig.6 Optical Motion Capture session

In order to perform the recorded animations, the template model is attached to an H-Anim LoA 2 [3] skeleton hierarchy. A proper skin attachment is essential for the skeletal deformation. The attachment is considered as assigning for each vertex of the mesh its affecting bones and corresponding weights. This method combines for each vertex the transformation matrix of the bones in accordance to their weight. Once the skin is properly attached to the skeleton, transformation of the bone automatically derives transformation of the skin mesh [6]. The captured animation from the models is finally applied to the H-Anim skeleton to obtain the animated bodies as shown in figure 6.

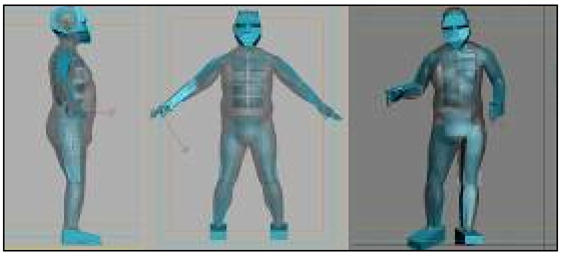


Fig.7 Optical Motion Capture session

3. Virtual Cloth Simulation

Fashionizer is a platform for 3D virtual prototyping of garments that has been developed in MIRALab, University of Geneva [8,9,10,11]. It integrates the design of 2D pattern, the simulation of fabrics, virtual human animation, comfortability evaluation, real-time animation & visualisation. The **Fashionizer** platform is based on general mechanical and collision detection schemes and allows for the simulation of multiple interacting garments and bodies. Combined to animation, Fashionizer is a powerful system allowing to compute the accurate behavior of cloth

- **Design of the patterns**

The workflow of Fashionizer takes an analogy to what actually happens during the creation of real clothing. Garments are composed of 2D patterns, describing the surface of a fabric. The dimension of these patterns corresponds to the size of the body to be dressed. In Fashionizer the patterns are built using an “editor” to specify 2D measurements. To work in 2D is more precise and easy to handle. The software is basically an editor of “Polygons” representing clothing in a number of 2D Polygons, connected by seam lines.

The creation of an entire garment is carried out in several stages:

First the creation of the pattern is made on a grid in the 2D window. The silhouette of the body appears in wireframe behind the pattern. The 3D window displays in real time the operations made in the 2D window and vice versa.

Therefore careful research on the clothing culture of the monks period is carried out and the collected data, such as patterns of different cloths, images and historical description , are used as base for the restitution of different types of clothes and their simulation around the bodies.

The cloths of the monks, the prior and subprior are the same. The monks wear a black ‘roccus’ over a white tunic. Both are made of wool (raw, not smooth). The tunic is a long dress that reaches to the feet. The ‘roccus’ is a long coat with wide sleeves and pointed cap.

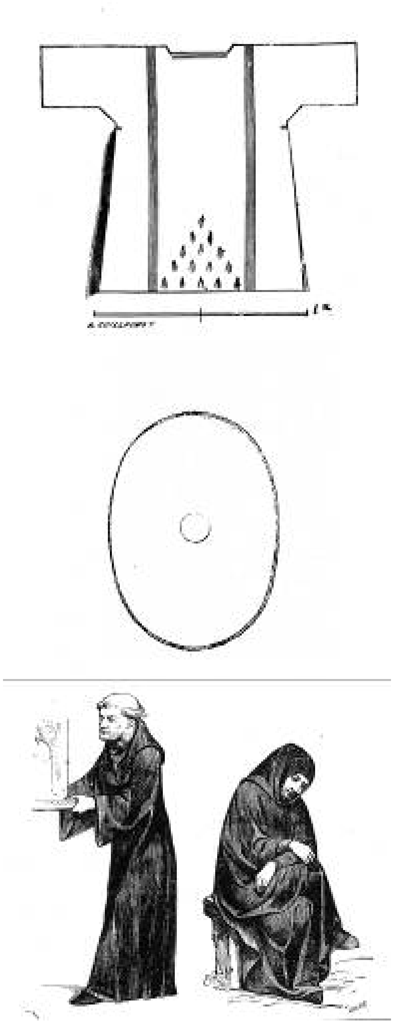


Fig.8 historical references

• Placement of the patterns around the body 3D

The patterns obtained in the preceding part are the placed around the virtual body and the seaming can be executed. The seams are forcing the patterns to approach and to pull the matching pattern borders together during simulation.

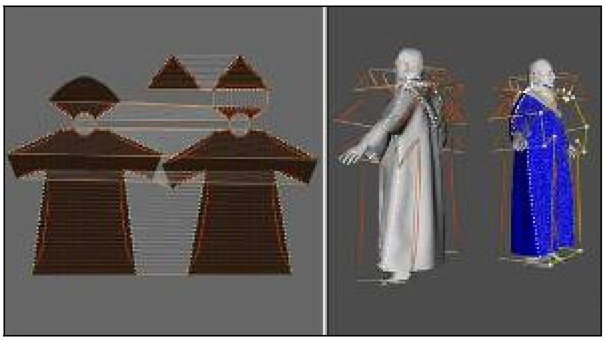


Fig.9 Placement of the 3D patterns

• Adding garment parameters

The garment parameters, and particularly the physical parameters of a texture (such as cotton, linen, silk, etc.), are adjusted through the garment material panel. It features two categories: environment (global parameters) and object (local parameters). Among the global simulation parameters, we will find gravity and finally collision distance and detection modes. Whereas the local parameters include elasticity, surface density, bending rigidity, friction values, Poisson coefficient, as well as viscosity and non-linear elasticity values, which are the mechanical properties of the garment.

• Fitting of the garment

Once the texturing and garment properties setting is done, the “fitting” of a garment can be started, by calling a mechanical simulation, which is forcing the surfaces to approach along the seam lines.The surface is deforming according to the shape of the body. The simulation engine first uses a simplified mechanical model, which is optimized in speed by leaving the physical parameters and environment parameters behind the calculation. The simulation has to be executed until the fabric is dynamically stabilized. The resulting position is a suitable starting point for the simulation of animation.

• Animation of clothing

The realistic clothing animation is simulated according to the movement of the virtual actor. This is possible thanks to the collision detection and friction with the surface of the body. The mechanical parameters are regulated at this stage and are adapted to the visual information. The adjustments of simulation parameters can be different from those used during the process of seaming and assembling of the garment. Mechanical simulation then gives the realism to the animation of clothing on the virtual character.

Figure.10 Exemples of Monks cloths simulation

References

[1] H. Seo, N. Magnenat Thalmann, "An Example-Based Approach to Human Body Manipulation", Graphical Models, Academic Press, Vol. 66, No. 1, pp. 1-23, January 2004

[2] Tecmath AG. Avilable from http://www.tecmath.com, Scanner operated by courtesy of ATC (Athens Technology Center S.A.).

[3] H-Anim specification. Avilable from http://ece.uwaterloo.ca/_H-ANIM/spec1.1.

[4] Bones Pro 2, Digimation. Avilable from http://www.digimation.com.

[5] H. Seo and N. Magnenat-Thalmann, An Automatic Modelling of Human Bodies from Sizing Parameters, in: Proceedings SIGGRAPH symposium on Interactive 3D Graphics, pp. 19–26, p. 234, 2003.

[6] Oxford Metrics, Real-Time Vicon8 Rt, 2000. Retrieved 10-04-01 from: http://www.metrics.co.uk/animation/realtime/realtimeframe.htm

[7] MotionAnalysis, Real-Time HiRES 3D Motion Capture System, 2000. Retrieved 17-04-01 from: http://www.motionanalysis.com/applications/movement/research/real3d.html

[8] P. Volino, N. Magnenat-Thalmann, "Implementing fast Cloth Simulation with Collision Response", Computer Graphics International Proceedings, IEEE Computer Society, pp 257-266, 2000.

[9] P. Volino, N. Magnenat-Thalmann, "Accurate Collision response on polygonal Meshes", Computer Animation 2000 Proceedings, IEEE Computer Society, 2000.

[10] P. Volino, N. Magnenat-Thalmann, "Comparing Efficiency of Integration Methods for Cloth Simulation", Computer Graphics International Proceedings, 2001.

[11] N. Magnenat-Thalmann, P. Volino, F. Cordier, "Avenues of Research in Dynamic Clothing", Computer Animation 2000 proceedings, IEEE Computer Society, pp 193-202, 2002.

[12] W.Lee, N.Magnenat-Thalmann, "Fast Head Modeling for Animation", Journal Image and Vision Computing, Volume 18, Number 4, pp.355-364, Elsevier, March, 2000.

[13] P. Volino, N. Magnenat-Thalmann, "Animating Complex Hairstyles in Real-Time", in: Conference VRST 2004.

EPOCH

Avatar-Based Interactive Storytelling: A New Cultural Heritage Technology

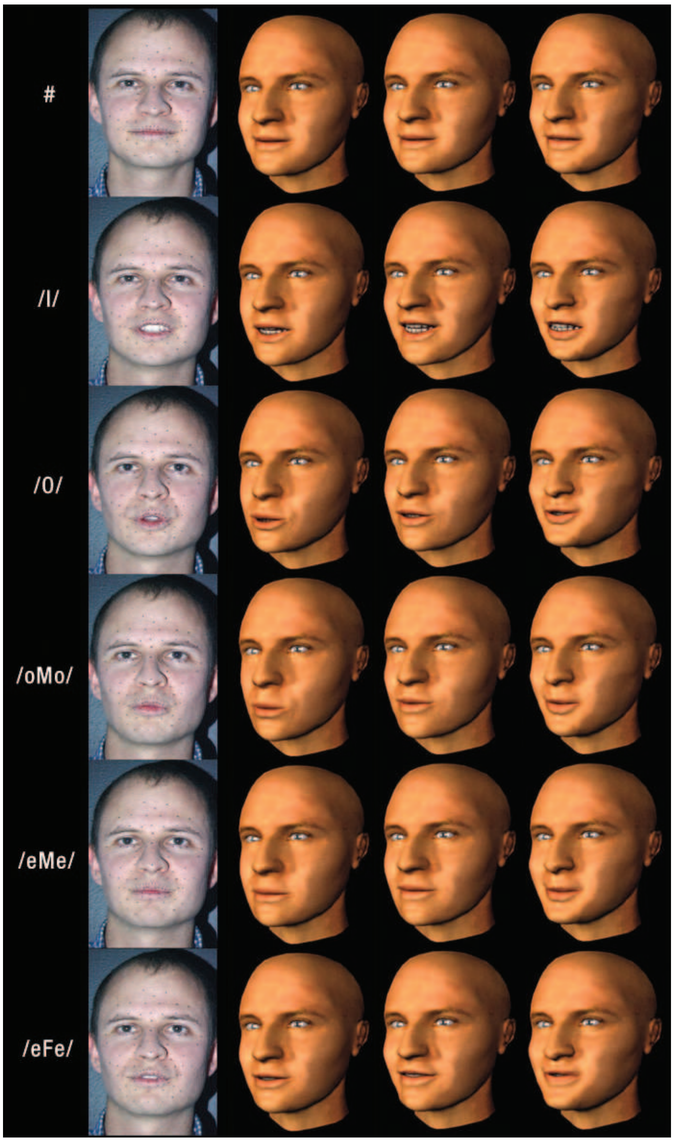
Stories About the Past Storytelling is a familiar and effective way to convey information to a general audience, especially in the field of Cultural Heritage. Yet with the increasing use of standardized interactive applications in museums and historic sites, it is difficult to appeal to a wide range of visitor interests. This problem can be solved through the use of interactive storytelling, which allows the visitor to choose from a large selection of subjects and themes to create a personalized "story", appropriate to his or her interests and the time available for the visit.

Presentation systems featuring interactive storytelling should be easy to update and maintain, without the expertise of technical specialists. Text-based interactive story telling applications, driven by a relatively simple XML or database application, allow museum and site staff to update the content by merely typing in texts.

Recorded narration is generally more effective than text in conveying information to the visitor, but its use in interactive storytelling has a major drawback concerning updating: changing minor parts of the content requires major investments in time and money (studio recording sessions, editing, digitizing) which often results in discouraging frequent updates and jeopardizing the entire update process. By adding synthetic voice capabilities to the interactive storytelling system, this problem of updating time and expense can be effectively addressed. Text-to-voice translation software makes it possible to create narration from the keyboard. Simple software tools can be used to alter intonation or stress significant words, making the synthetic speech more lifelike. The sound files can be easily exchanged in the application and they can provide narration in several languages.

A Virtual Guide An additional element can greatly enhance the effectiveness of an interactive storytelling system: the use of an avatar to serve as personal guide and storyteller for each visitor. 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. And this element too can be integrated into a system that is easily updated. The site of Akrotiri, Greece — one of the most important and most well-preserved sites of the Minoan Civilization — will be used as a test site for a showcase that demonstrates Avatar-Based Interactive Storytelling.

In the prototype system, the Avatar's speech and animation is fully synchronised with the synthetic voice, derived semi-automatically from written text in storytelling database. This integration can provide easily updatable multi-lingual, interactive explanations of major points of archaeological, cultural, and scientific interest in the form of self-authored "stories" conveyed by a virtual guide. The Akrotiri showcase will also utilize Virtual Humans to "populate" digital

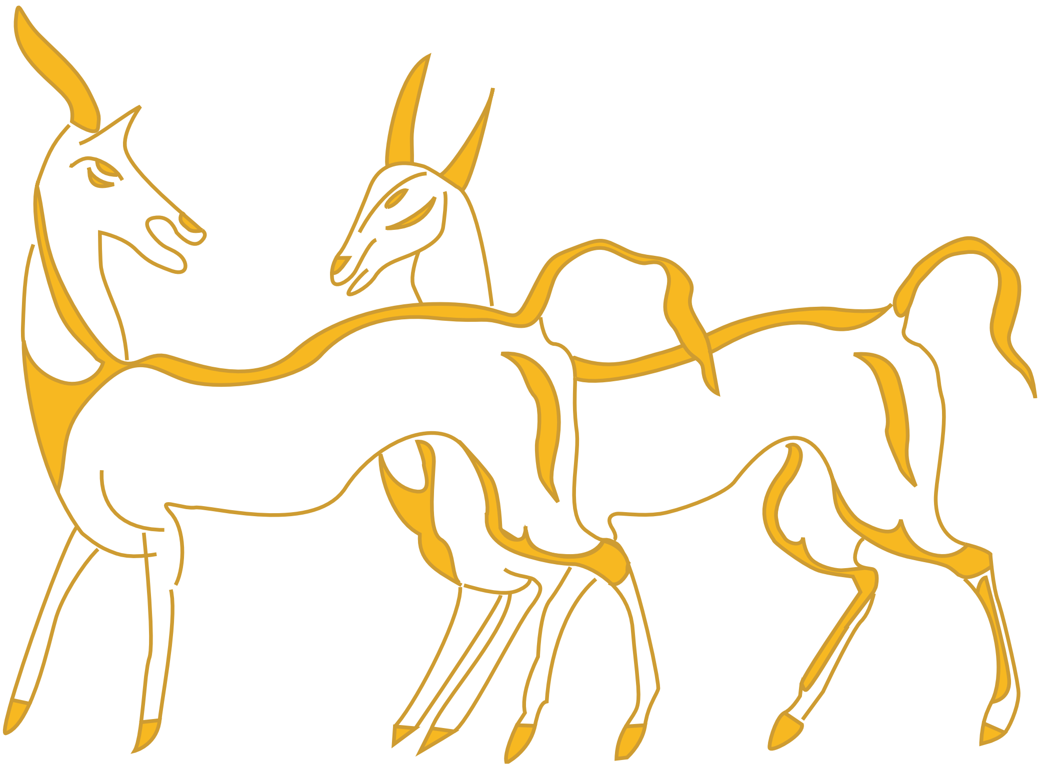


Digitalisation of a real person and his pronounciation of a set of phonemes yield a speaking avatar reconstructions of the site—and to address (through alternative versions) still unresolved questions about the ancient inhabitants of the site—their dress, their behavior, and their physical stature and appearance. Issues of alternative architectural reconstruction will also form a research priority for this showcase.

Technical specifications The showcase application is based upon the TimeScope 3 software, developed by the Ename Center and IBM, which is an Open Source XML and Java environment, build upon an Apache Tomcat server. Data entry for content specialists uses standardised forms to input texts and link images, animations and sounds. The text-to-speech software is provided by SVox, a spin-off company of ETH Zürich. The synchronised avatars are a development of the Computer Vision Lab of ETH Zürich, and use standardised emotions. The virtual humans are created and animated by MIRALab.

Partners This showcase is being realised, with the support of the Archaeological Society at Athens and Professor Christos Doumas, director of the excavations at Akrotiri, by:

- The Ename Center for Public Archaeology and Heritage Presentation, Belgium
- ETH Zürich, Computer Vision Laboratory, Switzerland
- MIRALab, University of Geneva, Switzerland



Interested?

Are you interested in this showcase? Do you think that this approach can help you in creating effective Cultural Heritage presentation projects or can be integrated in new research projects? Please contact Heidi Tency (heidi.tency@enamecenter.org) of the Ename Center at +32 55 232445.

EPOCH is a Network of Excellence on Intelligent Cultural Heritage within the IST (Information Society Technologies) section of the Sixth Framework Programme of the European Commission. EPOCH showcases demonstrate innovative solutions and technological integration for target application areas in the Cultural Heritage domain. As they are created with real world content, they stimulate creative thinking about the use of the technologies in Cultural Heritage, and are used to validate new technological approaches with key stakeholders in the Cultural Heritage domain. For more details, visit the project web site:

www.epoch-net.org

EPOCH is funded by the European Commission under the Community's Sixth Framework Programme, contract no. 507382. However, this leaflet reflects only the authors' views and the European Community is not liable for any use that may be made of the information contained herein.