

# Mobile Augmented Heritage: Enabling Human Life in ancient Pompeii

George Papagiannakis, Nadia Magnenat-Thalmann

MIRALab, University of Geneva, CUI

Battelle, Batiment A, 7 Route de Drize, CH-1227, Carouge, Switzerland

[papagiannakis@miralab.unige.ch](mailto:papagiannakis@miralab.unige.ch), [thalmann@miralab.unige.ch](mailto:thalmann@miralab.unige.ch)

## Abstract

*We propose a new methodology for real-time mobile mixed reality systems that feature realistic simulations of animated virtual human actors (clothes, body, skin, face) who augment real environments and re-enact staged storytelling dramas. Although initially targeted at Cultural Heritage Sites, the paradigm is by no means limited to such subjects. The abandonment of traditional concepts of static cultural artifacts or rigid geometrical and 2D textual augmentations with 3D, interactive, augmented historical character-based event representations in a mobile and wearable setup, is the main contribution of the described work as well as the proposed extensions to AR Enabling technologies: a VR/AR character simulation kernel framework with character to object interaction, a markerless camera tracker specialized for non-invasive geometrical registration on heritage sites and a PRT mixed reality illumination model for more consistent real-virtual real-time rendering. We demonstrate a real-time case study on the actual site of ancient Pompeii.*

## 1. INTRODUCTION

Mixed Realities [1] and their concept of cyber-real space interplay invoke such interactive digital narratives that promote new patterns of understanding. However, the "narrative" part, which refers to a set of events happening during a certain period of time and providing aesthetic, dramaturgical and emotional elements, objects and attitudes [8] is still an early topic of research. Mixing such aesthetic ambiences with virtual character augmentations [24] and adding dramatic tension has developed very recently these narrative patterns into an exciting new edutainment medium [8]. Since recently, AR Systems had various difficulties to manage such a time-travel in a fully interactive manner, due to hardware & software complexities in AR 'Enabling Technologies' [2]. Generally the setup of such systems was only operational in specific places (indoors-outdoors) or with specific objects which were used for training purposes rendering them not easily applicable in different sites. Furthermore, almost none of these systems feature full real-time virtual human simulation. With our approach, based on an efficient real-time tracking system, which require only a small pre-recorded sequence as a database, we can setup the AR experience with animated virtual humans anywhere, quickly. With the interplay of a modern real-time framework for integrated interactive virtual character simu-

lation, we can enhance the experience with full virtual character simulations. Even if the environmental conditions are drastically altered, thus causing problems for the real-time camera tracker, we can re-train the camera tracker to allow it to continue its operation. The proposed set of algorithms and methodologies aim to extend the “AR Enabling Technologies” in order to further support real-time, mobile, dramaturgical and behaviourised Mixed Reality simulations, as opposed to static annotations or rigid geometrical objects. Figure 1 depicts fully simulated virtual humans (skin, clothes, face, body) augmenting a cultural heritage site.

Figure 1. Example of mixed reality animated character acting a storytelling drama on the site of ancient Pompeii (view from the mobile AR-life system i-glasses)

## 1.1. Overview

As a preprocessing stage, our real-time markerless camera tracker system is being trained on the scene that is aimed to act as the mixed Reality stage for the Virtual actors. During real-time mobile operation and having already prepared the VR content for the virtual play, our system allows the user to be immersed fully in the augmented scene and for the first time witness storytelling experiences enacted by realistic virtual humans in mixed reality worlds. The minimal technical skills and hardware configuration required to use the system, which is based on portable wearable devices, allow for easy setup in various indoors and outdoors feature rich locations. Thus in Section 2 of this work we review the previous work performed in the main areas of “AR Enabling technologies”, such as camera tracking and illumination as well as the extensions that we propose: complete VR character simulation framework with character to object interactions as well as a new illumination model for MR. In Section 3 we present such a framework which is mandatory in order to handle the exponential complexity of virtual character drama, that traditional renderer-centric AR systems cannot anymore handle. In Section 4 we present our markerless camera tracker which allows for non-invasive camera tracking (no markers or tracking devices) with a robust boot-run solution i.e. that can recover from an incorrect registration due to user view on unrecognized parts of the scene. Thus visitors of the site can continue their experience with an instant ‘recovery’ of the camera tracker whereas other existing solutions would simply stop operation. A new illumination model and PRT method is presented in Section 5, in order to enhance the fidelity and consistency of ‘compositing’ virtual humans in AR. Finally in section 6 we present our two case studies in a controlled environment as well as on the site of ancient Pompeii and we epitomize with the discussion and conclusions sections 7 and 8 respectively.

## 2. RELATED WORK

On AR integrated platforms, a number of projects are currently exploring a variety of applications in different domains such as cultural heritage [3], training and maintenance [4] and games [5]. Special focus has recently been applied to system design and architecture in order to provide the various AR enabling technologies a framework [6] for proper collaboration and interplay. Azuma [2] describes an extensive bibliography on current state-of-the-art AR systems & frameworks. However, few of these systems take the modern approach that a realistic mixed reality application, rich in AR virtual character experiences, should be based on a complete VR Framework (featuring game-engine like components) with the addition of the “AR enabling Technologies” like a) Real-time Camera Tracking b) AR Displays and interfaces c) Registration and Calibration. Virtual characters were also used in the MR-Project [8] where a complete VR/AR framework for Mixed Reality applications had been created. Apart from the custom tracking/rendering modules a specialized video and see-through HMD has been devised. However, none of the aforementioned AR systems can achieve to date, realistic, complete virtual human simulation in AR featuring skeletal animation, skin deformation, facial-speech and clothes simulation. For realizing the dynamic notions of character based Augmented Heritage, the above features are a prerequisite.

Camera tracking methods can be broadly divided into outside-in and inside-out approaches, depending on whether the sensing device is located on the tracked object, or multiple sensing devices surround the tracked object. Technologies used to perform the tracking include mechanical [9], magnetic, optical, inertial, ultrasound or hybrids [10] of these. Due to the sensitive nature of the environment our tracking system had to function in, namely the ancient ruins of Pompeii, we opted to develop an inside-out optical system without the need for fiducial markers. Visual through-the-lens tracking is a widely researched topic and several papers have been published investigating different methods of achieving this goal. A software library has also been released called ARToolKit [11] which is widely used in the Augmented Reality community. This relies on large fiducials placed throughout the scene which are identified by the system and used to perform pose estimation. Davison [12] has published several papers based on Simultaneous Localization and Mapping (SLAM), a very promising real-time probabilistic visual method of tracking. Another popular approach is that reported by Fua [13] which is based on a prior scene model with real-time line and texture matching. The visual fiducial method commercialized by Radamec [14] has proven to be very accurate but has the disadvantage of requiring severe scene modifications and extensive setup.

High Fidelity real-time virtual character rendering methods have recently attracted new attention with variants of Precomputed Radiance Transfer (PRT) algorithms. [25] considers PRT on deformable characters but is based on precalculated, static spherical harmonic lightmaps and predefined vertex-based animations thus light cannot be rotated and the technique cannot scale to include BRDF simulation.

[26] allows for local PRT effects, based on texture-space bumps, wrinkles for deformable objects, but does not consider distant shadowing (e.g. limb to limb) such as it is typical in articulated characters or multi-segmented mesh hierarchies

[27] present an interesting PRT-variant method, but do not consider multi-segmented, deformable objects or dynamic lights.

[28] allows for soft-shadowing for dynamic, rigid body animated objects but considers only direct diffuse global illumination and high-frequency shadowing is calculated without rotations or in real-time (due to wavelets nature). Also it is unclear how this method could be applied to multi-segmented, deformable objects and with what performance penalty.

[29] recently presented a fast method for real-time soft shadows in dynamic scenes illuminated by large, low-frequency light sources. This work is the closest to our goals as it is applicable to virtual characters in real-time VR. However, it is unclear how transfer from multiple objects can be combined in the same hierarchy as in the case of multi-segmented virtual humans of the Humanoid Animation Group (Hanim) standard and introduced in a mobile AR framework (performance critical). Nevertheless this would be the most appropriate alternative to our approach on evaluating the hypotheses of high-fidelity illumination registration of virtual humans in Mixed Reality.

### **3. MR FRAMEWORK COMPONENTS FOR CHARACTER SIMULATION**

#### **3.1. AR-Life System design**

Our AR-Life system is based on the VHD++ [22], component-based framework engine developed by VRLAB-EPFL and MIRALab-UNIGE which allows quick prototyping of VR-AR applications featuring integrated real-time virtual character simulation technologies, depicted in [22]. The key innovation is focused in the area of component-based framework that allows the plug-and-play of different heterogeneous human simulation technologies such as: Real-time character rendering in AR (supporting real-virtual occlusions), real-time camera tracking, facial simulation and speech, body animation with skinning, 3D sound, cloth simulation and behavioral scripting of actions.

The integrated to the AR framework tracking component is based on a two-stage approach. Firstly the system uses a recorded sequence of the operating environment in order to train the recognition module. The recognition module contains a database with invariant feature descriptors for the entire scene. The runtime module then recognizes features in scenes by comparing them to entries in its scene database. By combining many of these recognized features it calculates the location of the camera and thus the user position and orientation in the operating environment. The main design principle was to maximize the flexibility while keeping excellent real-time performance. The different components may be grouped into the two following main categories:

- System kernel components responsible for the interactive real-time simulation initialization and execution.
- Interaction components driving external VR devices and providing various GUIs allowing for interactive scenario authoring, triggering and control.

Finally the content to be created and used by the system was specified, which may be classified into the two following main categories: a) Static and b) Dynamic content building blocks such as models of the 3D scenes, virtual humans, objects, animations, behaviors, speech, sounds, python scripts, etc.

### 3.2. MR registration and staging

Employing a markerless camera tracking solution for registering the CG camera according to the real one, is an added value advantage since it eliminates the use of external tracking devices or avoids polluting the real scene with the use of known fiducial markers. However, the issue that arises is how to geometrically calibrate the camera and define the scene fiducial origin in world coordinates. Especially as our MR scenes have animated virtual characters, initial character staging, scaling and orientation is a crucial factor in order to determine correct initial, life-sized, believable geometrical registration. In the pipeline described in section 4, boujou™ allows for an initial scene origin to be defined offline on a tracked scene feature. This feature though is not sufficient as for a number of characters and a storytelling scenario, designers would like to interactively direct, stage and adjust the action in real-time, according to their dramaturgical interest.

Therefore we propose a simple algorithm for determining the storytelling scene origin and orientation, harnessing the features of the underlying OpenGL scenegraph renderer camera metaphor and world node coordinates as depicted in Figure 2. We allow for interactive manipulation of the scene camera as well as separate scene global repositioning and scaling according to the following OpenGL formulas:

MODELVIEW MATRIX = [R | -R.t]

PROJECTION MATRIX =

$$\begin{bmatrix} \frac{2f}{w} & \frac{2s}{h} & -f - s + u_0 & 0 \\ 0 & \frac{2af}{h} & -af + v_0 & 0 \\ 0 & 0 & \frac{(p_f + p_n)}{(p_f - p_n)} & \frac{2p_f p_n}{(p_f - p_n)} \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

The intrinsic parameters of the camera are known in advance from boujou™ (Section 4), where  $f$  = focal length,  $s$  = axis skew,  $a$  = pixel ratio and  $(u_0, v_0)$  is the principal point. The ModelView and Projection matrices are used to set up the virtual camera metaphor and are provided in real-time for each tracked frame, by the underlying camera tracker (Section 4). Thus for interactive authoring-staging in the MR scene, two more controls are supplied: a) A single vector is mapped via keyboard controls on the translation part of the camera so that the camera can be furthermore tweaked within the tracked frame and b) A single 4x4 translation matrix is mapped as a virtual trackball metaphor, so that a designer can interactively stage the “Main Scene” scenegraph node that contains all the virtual augmentation. Since the basic renderer of the VHD++ framework is based on OpenScenegraph, all individual elements can have their transformation matrix modified. However, for the MR real-time authoring

stage, it is important that the whole staged experience can be initially positioned according to the real scene so that the camera tracking module can subsequently register it accordingly with the real scene.

Figure 2. Camera Coordinate Systems in MR

For the final geometrical registration, the following algorithm was employed, according to Figure 2 to modify the scenegraph:

1. Retrieve the camera image
2. Run the feature tracker on this image
3. Extract ModelView and Projection camera matrices
4. modify the combined camera matrix according to user authoring scaling/position controls (mapped virtual trackball mouse metaphor operation)
5. apply the combined/adjusted camera matrix to scenegraph renderer
6. move Occluder geometries as root nodes in the scenegraph with GL-DEPTH test set to OFF
7. set the background image acquired from the camera in a 2D projected screen textured quad (thus since background image applied as a texture it is hardware accelerated and independent of window resolution, as hardware extension for non power of 2 textures was employed)
8. modify the Main Scene root node according to user preference for further global positioning and scaling
9. execute draw threads on the whole scenegraph

Thus with the above simple and fast algorithm we were able to both stage our fiducial MR characters and scene Occluders without any performance drop in frame rate.

## **4. REAL TIME MARKERLESS CAMERA TRACKING ALGORITHM**

### **4.1. Introduction**

We have utilised a robust tracking system from 2D3™ which can recover from complete occlusion or extreme motion blur within one frame. This was achieved by making the system capable of booting at real-time speeds which makes it very robust and solves the problem of drift which is present with most sequential trackers. Our system is based on a two-stage approach. In the first stage we construct a model of the scene which consists of 3D coordinates together with invariant descriptors for the feature appearances and is described in section 4.2. The run-time stage described in section 4.3 uses this model to quickly and accurately locate the camera within the initialized scene.

### **4.2. Scene Modeling**

We construct our model of the scene using techniques well known in the Structure-from-Motion community as described by Zisserman [15] and Fitzgibbon [16]. The input to this procedure is a short movie sequence of typically 100 frames covering the scene to be tracked. We import this footage into a commercial program called *boujou* [17] which is used in the film industry to

extract structure and camera motion from motion sequences. The process is completely automatic and starts with tracking feature interest points throughout the sequence as shown by the ‘trails’ in Figure 3.

Figure 3 Feature interest points shown as trails(left) and 3D structure as dots and camera path in red (right)

The software then uses structure from motion techniques to automatically compute the scene geometry as well as the camera characteristics and motion. This is shown in Figure 3 with the structure as dots and the camera motion as a red line. This gives us a sparse model of the tracking volume.

For the tracking system to be able to instantly localize the camera we need an invariant representation of certain key features in the scene. Requirements for this descriptor are that it is scale-invariant, rotationally invariant and partially affine invariant. The SIFT features proposed by Lowe [18] offer us exactly this at relatively low computational cost.

The movie sequence of the scene is processed and all SIFT features are detected. We detect the most consistent and reliable features and use the scene structure information produced by *boujou* to assign a 3D coordinate to each. This list of 3D coordinates with SIFT descriptors then represents the scene model we use in real-time tracking.

### 4.3. Real-time Operation

We initialise the system by loading the scene database previously created which contains a matched list of invariant descriptors and 3D coordinates. The robustness of our system comes from its ability to initialize with a single frame as we will now describe. The first task for any given image is to locate points of interest by searching for local maxima and minima in a difference-of-Gaussian pyramid as shown in [18]. The SIFT descriptors are then calculated for each point using a local image gradient histogram. A typical image will result in approximately 300 detected features. Each of these should then be indexed into our scene database to find the nearest neighbour which is defined as the keypoint with the minimum Euclidean distance for the invariant descriptor.

Figure 4 Previously preprocessed scenes augmented with virtual objects in real-time. Note that the objects are still tracked even with the real scene partially occluded.

For a large database and with a few hundred features per image this euclidean distance becomes exorbitant to calculate. At such high dimensions the kd-tree algorithm offers no significant speed improvements over an exhaustive search. Beis 19 proposed using a modified kd-tree algorithm called best-bin-first (BBF) which is an approximate nearest-neighbour algorithm which can index a large number of features at very high speed.

Indexing all our detected features in the image gives us a set of known 3d reference points  $\mathbf{X} = [x_w, y_w, z_w, 1]$  in homogeneous world coordinates and their projections in homogeneous image coordinates  $\mathbf{x} = [x_i, y_i, 1]$  which enables us to solve for the projection matrix  $\mathbf{P}$

$$\mathbf{x} = \mathbf{P}\mathbf{X}$$

using the linear algorithm detailed by Ameller [20]. We only need 3 points to solve for the euclidean camera projection matrix since the intrinsics are available from *boujou* from the scene modelling stage.

The set of 3d reference points and images is likely to contain a varying number of false matches which will invalidate the results if used for the linear computation of the projection matrix. In order to make the system more robust we employ a variant of the RANSAC algorithm proposed by Nister [21] to discard these incorrect matches. This preemptive RANSAC algorithm runs in constant time which is desirable in a real-time system. The resultant projection matrix is the best linear fit to our data which we then improve by doing a nonlinear least squares optimization.

This projection matrix  $\mathbf{P} = \mathbf{K}\mathbf{R}[\mathbf{I} \ -\mathbf{t}]$  can be decomposed into the intrinsic matrix  $\mathbf{K}$ , the rotation matrix  $\mathbf{R}$  and the translation vector  $\mathbf{t}$ . This information is used by the rendering module to correctly overlay CG elements onto the video.

#### 4.4. Real-time camera tracking results

The tracking system was implemented as a multithreaded module which can take advantage of multiprocessor systems. The feature detection stage took between 30 and 90 ms on a Pentium IV 3.2 Ghz processor. This variation in processing time is due to an automatic level of detail controller which does less processing in highly detailed areas and more processing in sparse areas. The feature indexing and pose estimation took between 5 and 20 ms. This variation in processing time is due to the system detecting that sometimes very little motion occurred between subsequent frames and no linear pose estimation is necessary. In this case the system can use the previous position estimate and adjust it with a motion model. The above mentioned timings result in a frame rate of between 9 and 28 frames per second.

Figure 5 results tracks on the office scene(left) and on the maquette scene(right)

Figure 5 shows the features selected by the RANSAC algorithm as inliers as red crosses with their history as yellow lines. Figure 4(left) shows the same scene with virtual objects augmented. The system can handle moderate amounts of motion blur while still maintaining lock due to the multiscale feature detection algorithm used. Severe occlusion can also be handled without affecting the tracking capability as shown in Figure 4(right). If the user covers up the camera momentarily the system will obtain lock within 1 frame of viewing the known scene again. The registration error of virtual objects is consistently less than 1 pixel which demonstrates the accuracy of the approach used.

## 5. MR ILLUMINATION GPU BASED MODEL

The main contribution of this effort is to propose a new MR physically correct illumination model based on Precomputed Radiance Transfer (PRT) [26][29], that allows the handling of complex animated deformable virtual humans, and employing ‘real light’ information captured as High Dynamic Range (HDR) light probes to light in real time the simulated 3D elements. The application of such approach allows the implementation of a believable MR illumination registration between the real elements of the simulation and the virtual augmentations. It equally permits to achieve an interactive exposure matching between the real light, as perceived by the real AR camera, and the virtual exposure of the acquired, simulated Area Light which illuminates the augmented scene. In order to enhance the consistency between the illumination of the real scene and the virtual characters, we have chosen to extend and adapt the Precomputed Radiance Transfer (PRT) illumination model (Figure 6), to allow its application to the multi-segmented, deformable and animatable hierarchies that our animation system [30] is controlling. The main issue that we have encountered and solved, was to allow the multiple segments constituting the virtual skeleton of the 3D virtual humans, due to the presence of different joints, segments, cloth, hair meshes, to respond to a high dynamic range area light captured from the real scene with the method described in [30].

Since the final shading information is stored per vertex in the H-Anim virtual human hierarchy, it is independent from the underlying animation approach. Therefore, our virtual human animation system can easily be integrated within our MR framework to drive the animation of the skeleton of the simulated 3D deformable virtual humans. Moreover, it equally allows the inclusion of correct shading information and effect employing the blending schedule defined in Section 5.

Finally, in order to validate the results of our illumination registration in VR based on the extensions from [30], in this work we have compared our real-time rendered virtual characters with offline global illumination simulations of the same meshes. Various offline global illumination techniques were utilized such as Radiosity and Ambient Occlusion with area lights in order to quantify the visual differences phenomenologically. Figure 6 illustrates our comparisons.

Figure 6: Our extended precomputed radiance transfer for multi-segmented virtual actors illumination model (left) as opposed to offline Radiosity based (middle) and sky-light ray-tracing (right)

The offline rendering results have been computed utilizing the 3DSMax rendering engine and the same HDR light probe light was employed in all tests.

## 6. CASE STUDIES AND RESULTS

To meet the hardware requirements of this aim, a single Alienware P4 3.20 Ghz Area-51 Mobile Workstation was used, with a GeforceFX5600Go NVIDIA graphics card, an IEEE1394 Unibrain Camera for fast image acquisition in a video-see-through i-

glasses SVGAPro monoscopic HMD setup, for advanced immersive simulation. Our previous efforts were based on a client-server distributed model, based on two mobile workstations as at the beginning of our project a single laptop could not suffice in simulating in real-time such a complex MR application. To achieve the requirement of ‘true mobility’, a single mobile workstation is used currently in our final demonstrations as shown in the figure below. That was rendered possible only after the recent advances on the hardware side (processor, GPU) of mobile workstations and on our further software/algorithmic improvements (sections 3, 4, 5), in the streaming image capturing and introduction of hyper-threading and GPU calculations of the MR renderer.

Figure 7 AR-Life simulator system

In all our case studies we employed 5 fully simulated virtual humans [24], 20 smart interactive objects for them, 1 python script, 1 occluder geometry and in the case of the ‘lab maquette’ the 3D geometry of part of the thermopolium. The case studies statistics utilizing the above hardware configuration boil down to 20fps for the camera tracker and 17fps for the main MR simulation for the ‘lab maquette’ trial and 35fps and 25 fps respectively for the Pompeii trial.

## 6.1. Controlled environment setup trial

In order to validate that our integrated AR framework for virtual character simulation operates in different environments, we have tested the system directly in the ruins of Pompeii. However, in order to further continue AR tests in a controlled lab environment, a real paper ‘maquette’ was constructed (in both small-scale as well as large-scale) in order to resemble the actual Pompeii site that we visited for our first on site tests. This allowed us for extra fine tuning and improvement of our simulation and framework, without having to visit numerous times the actual site. Figure 8 depicts an example of augmenting the real small-scale ‘maquette’. In Figure 9 an example of augmenting the large-scale, ‘life-size’ maquette of the real Pompeiiian site is given, with the virtual Roman actor, rendered with our rendering approach.

Figure 8. Lab ‘maquette’ controlled AR tests. For optimum flexibility and tests the camera is detached from the HMD and moved freely within the ‘tracked area’ augmenting it with virtual characters (laptop monitor)

Figure 9. Our PRT-based illumination model for virtual humans in AR (right) as opposed to the standard Phong based model (left)

## 6.2. Pompeii and the thermopolium of Vetutius Placidus trial

With the help of the Superintendence of Pompeii [23], who provided us with all necessary archaeological and historical information, we have selected the ‘thermopolium’ (tavern) of Vetutius Placidus and we conducted our experiments there. The results are

depicted in the following Figure 10, **Error! Reference source not found.** and Figure 12 where the technologies employed for simulating and authoring our virtual humans were already described in [24].

Figure 10. The Real Pompeian ‘thermopolium’ that was augmented with virtual animated virtual characters.

Figure 11. The Mobile AR-Life simulator system (top left). The previously shown laptop is inserted in the backpack for the run phase.

Figure 12 Real-time virtual Pompeian characters in the real site of the Pompeiian thermopolium. Note the use of geometry ‘occluders’ that allow part of the real scene to occlude part of the virtual human

## 7. DISCUSSION AND FUTURE WORK

From the end-user point of view, the benefits of the described approach are significant. He/she can view a mobile superposition of the real world with a virtual scene representing fully simulated living characters in real-time. As we are able to occlude with real objects the virtual augmentations, we can therefore enforce the sensation of presence by generating believable behaviors and interaction between the real and virtual objects. However, someone could argue that no ‘parthenogenesis’ in the Mixed reality algorithmic field is exhibited by this study; our premise is that, as shown in Section 2, fully simulated virtual human actors have not yet appeared widely in real-time, HMD-Based Augmented Reality providing a complete, mobile, wearable, integrated methodology that can be re-applied on multiple AR contexts with a markerless tracked camera in real-time. Our key contribution has been into providing an exciting integrated methodology-application for mobile Augmented Heritage with integrated research and extensions in the ‘AR Enabling technologies’ 2. In [24] the initial steps of this methodology were described with some first interior AR tests. In the present work we complete this line of research with also outdoor AR tests as well as a new PRT-based illumination model for virtual human simulation.

Finally, there are a number of issues that need to be further improved as future work. For example the illumination registration issue is still not solved which will help for rendering more realistic the MR experience and we are currently experimenting with further camera compensation techniques. The real-time camera tracking performance can be further improved as well as the training process to be further shortened. Currently if the lighting conditions on the real scene are significantly altered, the system has to be retrained and a new database to be generated. In the virtual human simulation domain, an important aspect that we will be addressing in the future is interactivity between the real users and virtual actors in the AR domain, so that the virtual humans

can 'see' the real ones and 'feel' their presence. Finally, we would also like to employ different networked media, wireless networks and mobile devices (e.g. ultra mobile PCs) in order to provide a less heavy and more efficient mobile AR system and overall experience.

## 8. CONCLUSION

Nowadays, when laymen visit some cultural heritage site, generally, they cannot fully grasp the ancient vibrant life that used to be integrated in the present ancient ruins. This is particularly true with ruins such as the ancient city of Pompeii, where we would like to observe and understand the behaviors and social patterns of living people from ancient Roman times, superimposed in the natural environment of the city. With the extensions to "AR Enabling technologies" and algorithms that we propose for camera tracking, virtual human simulation and PRT-based AR illumination model, coupled under a complete real-time framework for character simulation, we aim provide new dramaturgical notions for Mixed Reality. Such notions could extend further research in MR and develop it as an exciting edutainment medium.

## Acknowledgments

The authors would like to thank Brian O'Kennedy and Andrew Stoddart for their work in real-time camera tracking. The work presented has been supported by the Swiss Federal Office for Education and Science and the EU IST Programme, in frame of the LIFEPLUS and INTERMEDIA EU projects.

## References

1. Milgram, P., Kishino, F., A Taxonomy of Mixed Reality Visual Displays, IEICE Trans. Information Systems, vol. E77-D, no. 12, pp. 1321-1329, 1994
2. Azuma, R., Baillot, Y., Behringer, R., Feiner, S., Julier, S., MacIntyre, B., Recent Advances in Augmented Reality, IEEE Computer Graphics and Applications, November/December 2001
3. Stricker, D., Dähne, P., Seibert, F., Christou, I., Almeida, L., Ioannidis, N., Design and Development Issues for AR-CHEOGUIDE: An Augmented Reality-based Cultural Heritage On-site Guide, EuroImage ICAV 3D Conference in Augmented Virtual Environments and Three-dimensional Imaging, Mykonos, Greece, 30 May-01 June 2001
4. Wohlgemuth, W., Triebfürst, G., ARVIKA: augmented reality for development, production and service, DARE 2000 on Designing augmented reality environments, 2000

5. Thomas, B., Close, B., Donoghue, J., Squires, J., De Bondi, P., Morris, M., Piekarski, W., ARQuake: An Outdoor/Indoor Augmented Reality First Person Application, 4th Int'l Symposium on Wearable Computers, pp 139-146, 2000
6. Gamma, E., Helm, R., Johnson, R., Vlissides, J., Design Patterns: Elements of Reusable Object-Oriented Software, Addison-Wesley, 1994
7. Billinghurst, M., Poupyrev, I., Shared Space: Mixed Reality Interface for Collaborative Computing, Imagina'2000 Official Guide: Innovation Village Exhibition, , pp. 52, 2000
8. Tamura, H., Yamamoto, H., Katayama, A., Mixed reality: Future dreams seen at the border between real and virtual worlds, Computer Graphics and Applications, vol.21, no.6, pp.64-70. 2001
9. Mark Roberts Motion Control, Encoded Cranes and Dollies, <http://www.mrmoco.com>, 2004
10. Foxlin, E., Harrington, M., Pfeifer, G., Constellation: A wide-range wireless motion-tracking system for augmented reality and virtual set applications, ACM SIGGRAPH '98, pp. 371-378, 1998
11. ARToolKit, <http://www.hitl.washington.edu/artoolkit>, 2007
12. Davison, A.J., Kita, N., 3D simultaneous localisation and map-building using active vision for a robot moving on undulating terrain, IEEE Conference on Computer Vision and Pattern Recognition, 2001
13. Vachetti, L., Lepetit, V., Fua, P., Combining Edge and Texture Information for Real-Time Accurate 3D Camera Tracking, International Symposium on Mixed and Augmented Reality, Arlington, VA, 2004
14. Thomas, G.A., Jin, J., Niblett, T., Urquhart, C., A Versatile Camera Position Measurement System For Virtual Reality TV Production, International Broadcasting Convention, Amsterdam, pp. 284-289, 1997
15. Hartley, R.I., Zisserman, A., Multiple View Geometry in Computer Vision, Second Edition, Cambridge University Press, 2004
16. Fitzgibbon, A.W., Zisserman, A., Automatic Camera Recovery for Closed or Open Image Sequences, ECCV, pp. 311-326, 1998
17. Boujou, <http://www.2d3.com>, 2007

18. Lowe, D., Distinctive Image Features from Scale-Invariant Keypoints, International Journal of Computer Vision, 60, 2 (2004), pp. 91-110
19. Beis, J., Lowe, D., Shape Indexing using approximate nearest-neighbour search in high dimensional spaces, Computer Vision and Pattern Recognition, Puerto Rico, pp. 1000-1006, 1997
20. Ameller, M., Triggs, B., Quan, L., Camera Pose Revisited – New Linear Algorithms, ECCV, 2000
21. Nister, D., Preemptive RANSAC for Live Structure and Motion Estimation, ICCV Volume 1, pp. 199, 2003
22. Ponder, M., Papagiannakis, G., Molet, T., Magnenat-Thalmann, N., Thalmann, D., VHD++ Development Framework: Towards Extendible, Component Based VR/AR Simulation Engine Featuring Advanced Virtual Character Technologies, IEEE Computer Society Press, CGI Proceedings, pp. 96-104, 2003
23. Archaeological Superintendence of Pompeii, <http://www.pompeisites.org>, 2004
24. Papagiannakis, G., Schertenleib, S., O’Kennedy, B., Poizat, M., Magnenat-Thalmann, N., Stoddart, A., Thalmann, D., Mixing Virtual and Real scenes in the site of ancient Pompeii, Journal of Computer Animation and Virtual Worlds, pp. 11-24, Volume 16, Issue 1, John Wiley and Sons Ltd, February 2005
25. Xu, R., Pattanaik, S., Hughes, C. E., Real-time Rendering of Dynamic Objects in Dynamic, Low-frequency Lighting Environments, Proc. of Computer Animation and Social Agents04, 2004
26. Sloan, P.P., Luna, B., Snyder, J., Local, Deformable Precomputed Radiance Transfer, Proc. of ACM SIGGRAPH05, ACM Press, 2005
27. Tamura, N., Johan, H., Insita, T., Deferred shadowing for real-time rendering of dynamic scenes under environment illumination, Computer Animation and Virtual Worlds, vol. 16, pp. 475-486, 2005
28. Zhou, K., Hu, Y., Lin, S., Guo, B., Shum, H., Precomputed Shadow Fields for Dynamic Scenes, Proc. of ACM SIGGRAPH05, ACM Press, 2005
29. Ren, Z., Wang, R., Snyder, J., Zhou, K., Liu, X., Sun, B., Sloan, P.-P., Bao, H., Peng, Q., Guo, B., Real-time Soft Shadows in Dynamic Scenes using Spherical Harmonic Exponentiation, Proc. of ACM SIGGRAPH06, ACM Press, 2006

30. Egges, A., Papagiannakis, G., Magnenat-Thalmann, N., Presence and Interaction in Mixed Reality, The Visual Computer, Springer-Verlag (accepted), 2007

## List of Figures



### Augmented Life: first experiment on site



*Virtual scenes calculated by the computer*



*What you see in the 3D glasses: virtual scenes + real decor*

MIRALab - University of Geneva

Figure 1. Example of mixed reality animated character acting a storytelling drama on the site of ancient Pompeii (view from the mobile AR-life system i-glasses)

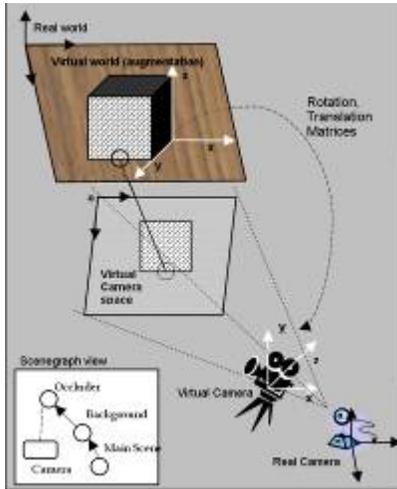


Figure 2. Camera Coordinate Systems in MR



Figure 3 Feature interest points shown as trails(left) and 3D structure as dots and camera path in red (right)



Figure 4 Previously preprocessed scenes augmented with virtual objects in real-time. Note that the objects are still tracked even with the real scene partially occluded.

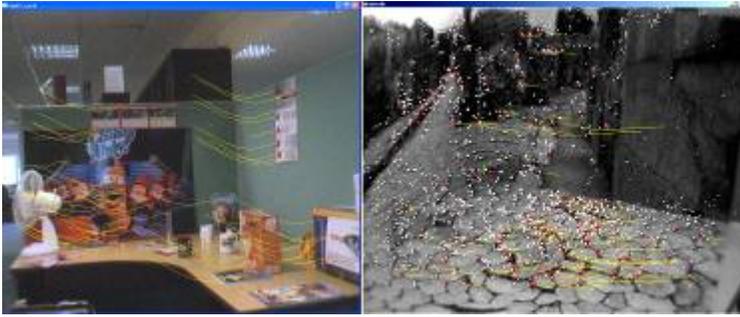


Figure 5 results tracks on the office scene(left) and on the maquette scene(right)



Figure 6 Our extended precomputed radiance transfer for multi-segmented virtual actors illumination model (left) as opposed to offline Radiosity based (middle) and sky-light ray-tracing (right)

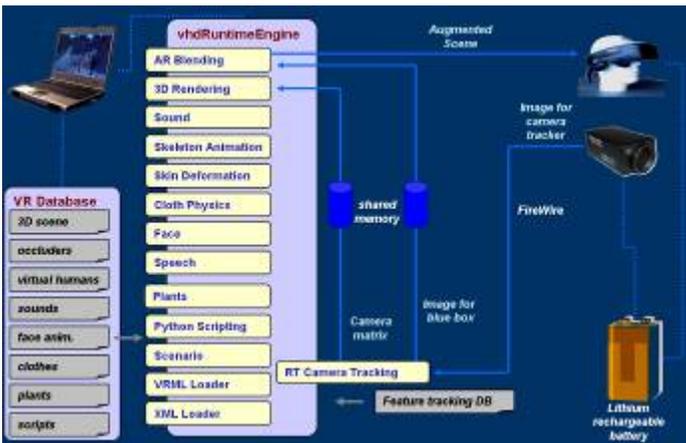


Figure 7 AR-Life simulator system



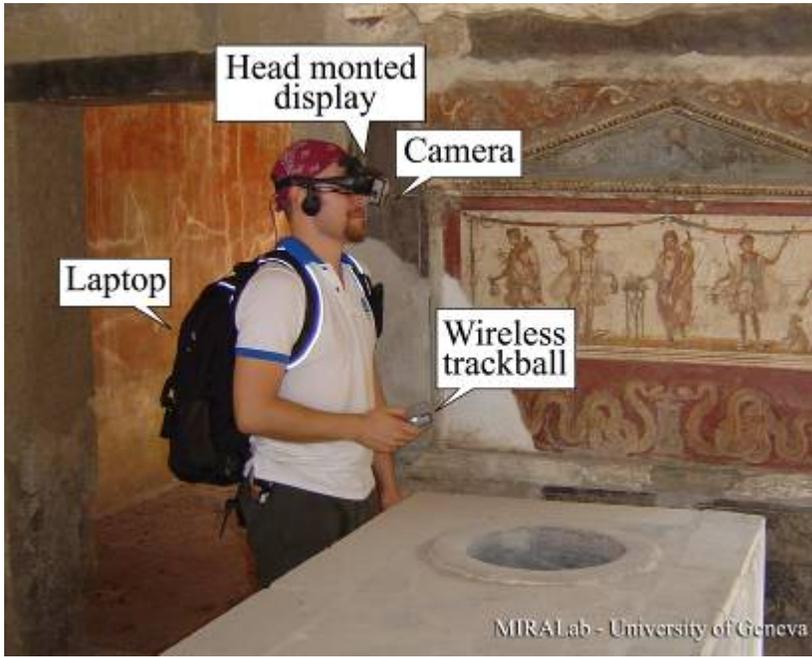
Figure 8 Lab ‘maquette’ controlled AR tests. For optimum flexibility and tests the camera is detached from the HMD and moved freely within the ‘tracked area’ augmenting it with virtual characters (laptop monitor)



Figure 9 Our PRT-based illumination model for virtual humans in AR (right) as opposed to the standard Phong based model (left)



Figure 10 The Real Pompeian ‘thermopolium’ that was augmented with virtual animated virtual characters.



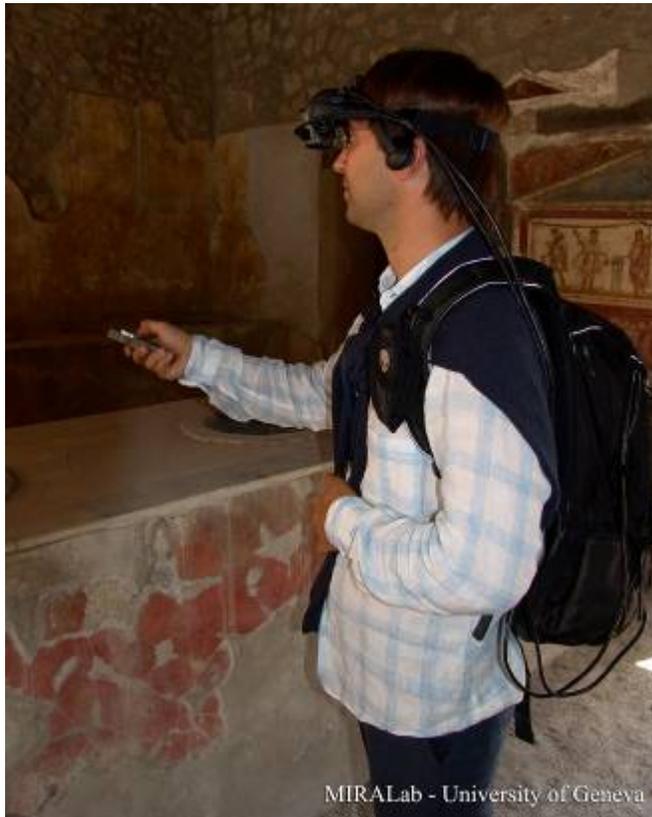


Figure 11 The Mobile AR-Life simulator system (top left). The previously shown laptop is inserted in the backpack for the run phase.

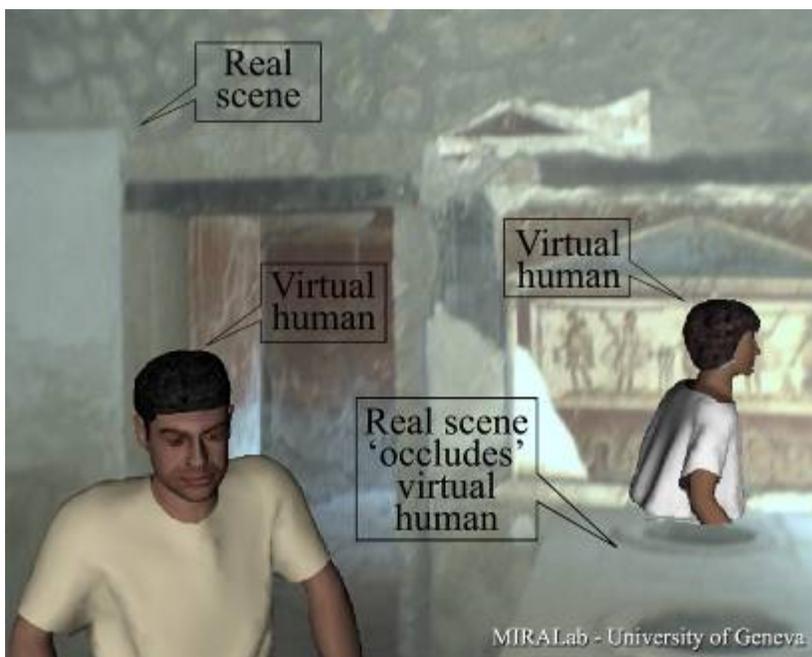


Figure 12 Real-time virtual Pompeian characters in the real site of the Pompeian thermopolium. Note the use of geometry 'occluders' that allow part of the real scene to occlude part of the virtual human