A

APPLICATIONS OF INTERACTIVE VIRTUAL HUMANS IN MOBILE AUGMENTED REALITY

Nadia Magnenat-Thalmann, George Papagiannakis, Parag Chaudhuri
MIRALab, University of Geneva, Switzerland

Definition: *Virtual humans are used as interfaces as well as real-time augmentations (three-dimensional computer-generated superimpositions) in real environments, as experienced by users though specialized equipment for enhanced mobility (e.g. ultra mobile PCs and video see-through glasses)*

Recent advances in hardware and software for mobile computing have enabled a new breed of mobile Augmented Reality systems and applications featuring interactive virtual characters. This has resulted from the convergence of the tremendous progress in mobile computer graphics and mobile AR interfaces. In this paper, we focus on the evolution of our algorithms and their integration towards improving the presence and interactivity of virtual characters in real and virtual environments, as we realize the transition from mobile workstations to ultra-mobile PC's. We examine in detail three crucial parts of such systems: user-tracked interaction; real-time, automatic, adaptable animation of virtual characters and deformable pre-computed radiance transfer illumination for virtual characters. We examine our efforts to enhance the sense of presence for the user, while maintaining the quality of animation and interactivity as we scale and deploy our AR framework in a variety of platforms. We examine different AR virtual human enhanced scenarios under the different mobile devices that illustrate the interplay and applications of our methods.

Introduction

Mixed Reality [1] has been depicted as a continuum that includes both Virtual Reality (VR) as well as Augmented Reality (AR). Traditionally the rich content needed for the complex, immersive simulations of VR dictated a desktop hardware setup, whereas the 2D or static 3D superimpositions on a real scene allowed for mobile, wearable (albeit cumbersome) systems. As the expectations and applications of AR have increased with the recent performance boost of mobile graphics on mobile workstations, modern mobile AR simulations have reached unparalleled levels of complexity, featuring advanced 3D simulations with animations, deformations and more realistic real-time rendering. However, these effects were achieved at the expense of the mobility of AR systems, as they were based on combinations of mobile workstations. Recently a new class of mobile devices has arrived, the Ultra Mobile PC (UMPC) that includes similar hardware and software capabilities of mobile workstations. Moreover, PDAs have also been merged with mobile phones, allowing for new opportunities based on their advanced hardware as well as more programming friendly software environments, operating systems and APIs. In this work we summarize our research efforts for the last 5 years, where advanced real-time 3D augmentations of fully simulated virtual characters (body, face and cloth simulation) were brought into mobile AR. In order to achieve such simulations in real-time, we initially commenced with a set-up of two mobile workstations, then migrated to a single workstation. Very recently our efforts have resulted in adapting our interactive virtual character simulations to a UMPC. This progress toward a more mobile AR system is depicted in Figure 1. It is interesting to note, however, that it is based on the same component-based 3D simulation framework [7].

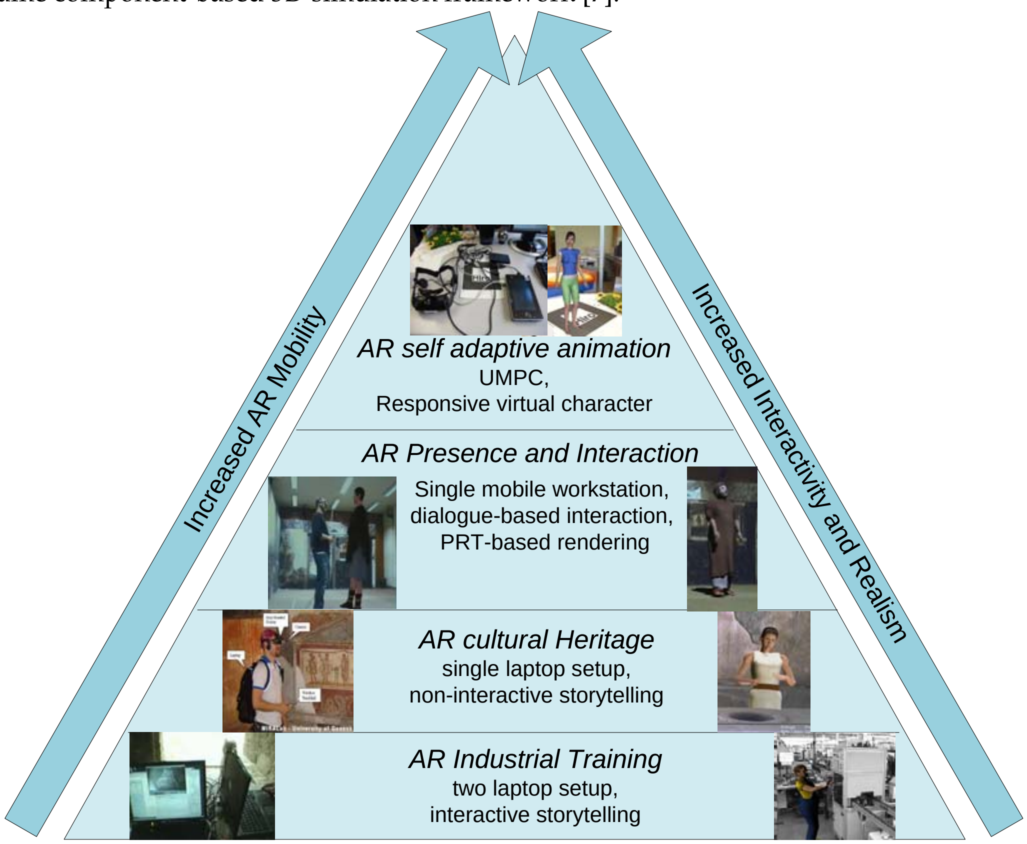

Figure 1. Evolution of interactive virtual humans and their presence, interaction and animation on mobile AR systems

Mobile Workstations & UMPCs for mobile AR

A number of systems [3] have employed mobile workstations (high-end laptops), often aggregated together with other mobile equipment, often carried in a backpack (weighting 1-6 kg), so that the user can freely move in the real environment and have their hands free for input and interaction. These backpacks include amongst others, mobile workstations such as Dell™ Inspiron & Precision, Alienware™ and JVC™ sub-notebooks. Although severe ergonomic issues are apparent due to the size and weight of the backpack, it allows researchers to focus on their research without the constraints that smaller devices often present, namely computational power, operating system and hardware connectivity. Almost all of the desktop computing system can be made mobile by using high-end notebook computers. However, due to the backpack setup, the use of head mounted displays (HMDs) is enforced as opposed to handheld display that other devices can offer. The next step towards this direction is the employment of ultra mobile PCs (UMPCs), that could provide both handheld as well as HMD viewing capabilities. The usage of ultra mobile PCs is a very recent trend in mobile AR systems is. UMPCs are based on the Microsoft Origami™ specification released in 2006 and have been developed jointly by Microsoft™, Intel™, and Samsung™, among others. UMPCs are basically small-factor mobile PCs capable of running Microsoft Windows XP™ or Vista™. A number of researchers have started employing them in AR simulations [2].

The VHD++ AR/VR mobile framework

AR systems rely heavily on the interplay of heterogeneous technologies. Because of that interdisciplinary character, the AR domain can be viewed as a melting pot of various technologies, which although complementary, are non-trivial to put together. The VHD++ framework [7] is a software development framework that supports production of high performance, real-time, interactive, audio-visual applications. Traditionally it had a core composed of 3D graphics, 3D sound and advanced synthetic character simulation but recently many other technologies like networking, database access, artificial intelligence, content creation and diagnostics have been added. Instead of conventional development the applications are being created by reuse and customization of the existing, fully operational design and code. For more details the reader is invite to look at [7], [8]. In the following sections we present four works that show the evolution of our AR technologies as we move toward a more mature mobile AR platform. We will see how these helped us improve our hardware and software development platform significantly as we consistently moved toward more believable and rich AR experiences on mobile platforms.

Instructive virtual industrial trainer based on dual mobile workstations

In this work, a novice user is trained to use complex machines by a virtual teacher showing them how to correctly operate machinery. Including real machinery and surroundings into the interactive simulation increases realism and decreases time that is required to build complex virtual environments and the computation cost involved in rendering them. Figure 2 illustrates this approach: a virtual worker demonstrates how to use a machine in a complex industrial environment.

Figure 2. A virtual human demonstrating the use of a real machine

The user of the AR system can change the point of view at any time, and since the camera position is correctly registered to the real scene, the virtual worker is always correctly blended into the streaming video. We therefore make a twofold contribution:
– An accurate real-time vision-based camera tracker, which is responsible for the registration of the virtual humans into streaming video and does not require any engineering of the environment.
– It's integration into an existing VR system, called VHD++ that provides the interface with the user and the rendering of realistic virtual humans.

Since VHD++ is a modular component-based framework, the integration of the tracking part was done as a plug-in and was straightforward. The VR part that has been used in this application integrates technologies, such as real-time 3D rendering, skeleton and skin animation and behavioral control together. VHD++ virtual humans show large range of animation capabilities. In our case we used keyframe animation of the virtual human body. For some general movement, such as walking, pointing and grabbing, an inverse kinematics module can also be used. The framework aims to act as a real-time, extensible audiovisual framework with dedicated support to VR/AR real-time virtual characters.

Mobile AR cultural heritage guide on a single laptop

Figure 3. A single laptop employed in a backpack for autonomous mobile AR.

In this application, we migrate from a configuration of two connected laptops to a single laptop configuration. This work is centred on the innovative revival of life in ancient frescos-paintings in ancient Pompeii and creation of narrative spaces [5]. The revival is based on real scenes captured on live video augmented with real-time autonomous groups of 3D virtual fauna and flora (Figure 3). The metaphor, which inspires the project approach, is oriented to make the "transportation in fictional and historical spaces", as depicted by frescos/paintings, as realistic, immersive and interactive as possible. Thus the ancient characters of the frescos/paintings (including humans and plants) will be revived and simulated in real-time 3D, exhibiting in a new innovative manner their unique aesthetic, dramaturgical and emotional elements. The whole experience is presented to the user on-site in Pompeii during their visit, through an immersive, mobile Augmented Reality-based Guide. This AR platform is also based on the VHD++ component-based framework. The various technologies used in this work include a plug-and-play combination of different heterogeneous technologies such as: Real-time character rendering in AR, real-time markerless camera tracking, facial simulation and speech, body animation with skinning, 3D sound, real-time cloth simulation and behavioural scripting of actions. To meet the hardware requirements of this aim, a single Dell Pentium 4 M50 Mobile Workstation was used, with a Quadro 4 500 GL Nvidia graphics card, a firewire Unibrain camera or USB Logitech web camera for fast image acquisition send on a video-see-through TekGear monoscopic HMD, for an advanced and immersive simulation. We started, as before, on a client-server distributed model, based on two mobile workstations. However, to achieve the requirement of 'true mobility', we migrated to a single mobile workstation. This is now used in our current demonstrations, after improvements in the streaming image capturing and introduction of hyper-threading in the platform code. We based our system on a real-time markerless camera tracking method from 2d3™ where the integrated camera tracker is able to self-initialize anywhere within the tracking environment without any intervention from the user as well as recover immediately in case of degenerate tracking (i.e., looking out of the designated area). In effect this means that instead of calculating relative changes in rotation and translation, we calculate absolute rotation and translation for every frame. This has the advantage of avoiding the problem of drift, and also ensures instant recovery after tracking was lost due to excessive motion blur or occlusion. The basic algorithm is based on "structure from motion" techniques and described more in detail in [8].

Interactive, dialogue based and advanced rendered virtual characters on a single mobile workstation

In this work, the previous approaches are extended to allow for interaction, animation and global illumination of virtual humans for an integrated and enhanced presence in AR. The interaction system comprises of a dialogue module that is interfaced with a speech recognition and synthesis system. In addition to speech output, the dialogue system also generates face and body motions, which are in turn passed on to the virtual human animation layer. All these different motions are generated and blended online, resulting in a flexible and realistic animation. Our robust rendering method operates in accordance with this animation layer and is based on an extension for dynamic virtual humans. The extended Precomputed Radiance Transfer (PRT) illumination model used results in a realistic display of such interactive virtual characters in complex mixed reality environments. The presented scenario illustrates the interplay and application of our methods, glued under the VHD++ framework for presence and interaction in mixed reality. It features a real human that engages in conversation with a virtual one in AR and witnesses the virtual human, correctly registered and aligned in natural size (human scale) in the real environment, based on the markerless camera tracker from [8] (see Figure 4).

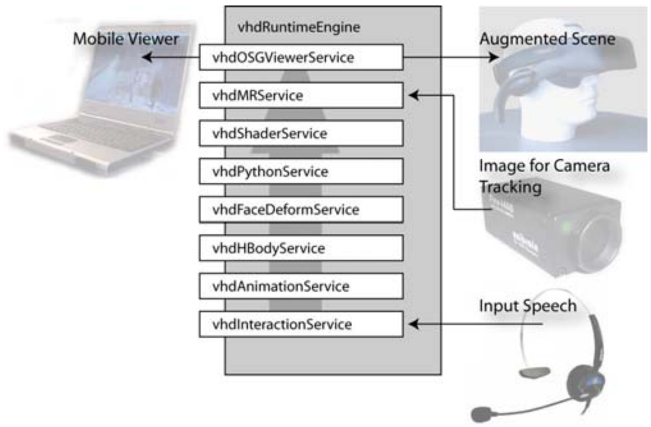

Figure 4. A single mobile workstation for advanced presence, PRT-based rendering and dialogue based interaction in AR

The software architecture is extended here for maintenance of the consistent simulation and interactive scenario state that can be modified with python scripts at run-time. To keep good performance, the system utilizes five threads. One thread is used to manage the updates of all the services that we need to compute, such as human animation, cloth simulation or voice (sound) management. A second thread is used for the 3D renderer that obtains information from the current scenegraph about the objects that must be drawn, in addition to the image received from the camera. It changes the modelview matrix accordingly to the value provided by the tracker. The third thread has the responsibility of capturing and tracking images. The fourth thread manages the update process of the interaction system, by parsing the script to see whether or not an action needs to be taken by the virtual human. The last thread is the python interpreter, which allows us to create scripts for manipulating our application at the system level, such as generating scenario-based scripted behaviors for the human actions (key-frame animation, voice, navigation combined to form virtual short plays). The AR system presented in Figure 4 features immersive, real-time, interactive simulation supplied with proper information during the course of the simulation. This, however, requires the components to be very diversified and thus their development is an extremely laborious process involving long and complex data processing pipelines, multiple recording technologies, various design tools and custom made software. The various 3D models to be included in the virtual environments like virtual humans or auxiliary objects have to be created manually by 3D digital artists. The creation of virtual humans require to record motion captured data for realistic skeletal animations as well as a database of gestures for controlling face and body animations. Sound environments, including voice acting, need to be recorded in advance based on storyboards. For each particular scenario, dedicated system configuration data specifying system operational parameters, parameters of the physical environment and parameters of the VR devices used have to be defined as well as scripts defining behaviours of simulation elements, in particular virtual humans, have to be created. These scripts can modify any data in use by the current simulation in real-time.

AR self adaptive animations on a UMPC

We explained in the previous sections that creating more and more believable and interactive characters becomes an increasingly laborious process. To alleviate this problem, we present in this section, our recent work on a simple and fast method to author self adaptive character animations that respond automatically to changes in the user's perspective or point of view in real-time, suitable for simulation in a UMPC. The animator creates a set of example key animations for the characters assuming the user is viewing the animation from different key viewpoints in the world. When the user actually interacts with the character, the user's actual point of view is tracked in real-time by using computer vision techniques or by simple user controlled input methods. The tracked position of the user's viewpoint, with respect to the key viewpoints, is then used to blend the example key animations, in real-time (see Figure 5). Thus, the animation of the character adapts itself in response to the changes in the user's viewpoint [10]. We demonstrate that our method is simpler and more efficient than other techniques that can be used to obtain similar results. We also show a working, prototype implementation of our method with a simple example in mobile Augmented Reality on a UMPC. The basic aim of this work, in comparison to the previous approaches is to:

- Allow for the same AR framework to be utilised effectively on a mobile PC as well as on a UMPC
- Allow for virtual characters to react to users presence during an interactive session, a shortcoming of previous methods

Due to the limited graphics acceleration of the UMPC, allow for 3D content adaptation of the same 3D augmentation. For e.g. The UMPC used did not support OpenGL 2.0 or the OpenGL Texture Rectangle extensions. Thus, both virtual character animation and skinning had to be calculated without accelerated vertex buffer objects and the video see-through HMD camera grabbing had to utilize simple Texture2D objects.

Implementation

A character pose is a hierarchical tree of rigid transformations. If we linearize this tree by doing a fixed traversal on it, we get a list of transformations. Rigid transformations can be represented as unit dual quaternions [10].

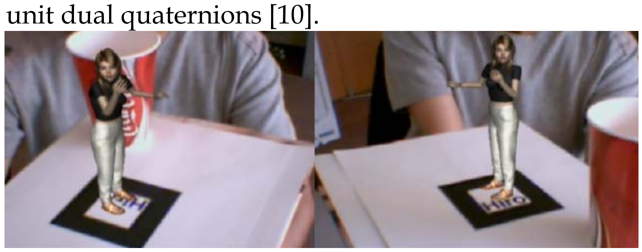

Figure 5. Self-adaptive animation based on AR user perspective

We define a character pose as a list of dual quaternions. An animation is only a time varying sequence of poses or by extension, a time varying list of dual quaternions. An animator creates a set of key animations that represent the way the character should react when the user approaches or looks at the character from different directions. We refer to these directions as key viewpoints or key cameras. Now the user's point of view is tracked in real-time by using known camera tracking algorithms. In this example, we have used ARToolkit for this purpose. We recover the current transform for a tracked marker and place our character on that position. The pose of the marker is used to infer the position of the current camera looking at the scene. We can also use our previous markerless camera tracking method for this. The current tracked camera is used to compute a weighted blend of the key animations to get the current animations that the user can see. The weights are computed on the basis of the position of the current camera in the space of key cameras defined earlier.

Figure 6. A wearable UMPC with the i-glasses HMD, battery pack and usb webcam.

Recently, UMPCs provide unique opportunities for mobile applications in terms of code portability as well as performance. Due to performance capabilities (CPU and GPU) several allowances have to be made and content has to be adapted to better fit the mobile experience that the UMPC offers. As we have recently witnessed the merging of PDAs with phones, it could be possible in the near future to witness a further merge between UMPCs with mobile phones.

Conclusions and Acknowledgements

Here we have presented an overview of our research work that has been carried out over the last years on presence, interaction and animation issues of 3D virtual characters in mobile AR systems. This effort is a material witness to the evolution and progress of state of the art in such systems, bringing richer and more believable content within the grasp of modern day mobile AR systems, networked media and computer graphics based 3D simulations. The currently presented work has been supported by the INTERMEDIA 38419 EU Project in frame of the EU IST FP6 Programme.

See: Virtual and Augmented Reality

References

[1] Azuma, R., Baillot, Y., Behringer, R., Feiner, S., Julier, S., MacIntyre, B., “Recent Advances in Augmented Reality”, IEEE Computer Graphics and Applications, 21(6):34-47, 2001.

[2] Barakonyi, I., Schmalstieg, D., “Ubiquitous Animated Agents for Augmented Reality”, Proceedings of ISMAR 2006 (IEEE and ACM International Symposium on Mixed and Augmented Reality), pp. 145-154, 2006.

[3] Papagiannakis, G., Singh, G., Magnenat-Thalmann, N., "A survey of mobile and wireless technologies for augmented reality systems", Journal of Computer Animation and Virtual Worlds, John Wiley and Sons Ltd, 2008, (to appear).

[4] Vacchetti, L., Lepetit, V., Ponder, M., Papagiannakis, G., Fua, P., Thalmann, D., Magnenat-Thalmann, N., "Stable Real-time AR Framework for Training and Planning in Industrial Environments", Virtual Reality and Augmented Reality Applications in Manufacturing, Ong, Soh K., Nee, ISBN:1-85233-796-6, Springer-Verlag, 2004.

[5] Papagiannakis, G., Schertenleib, S., O'Kennedy, B., Arevalo-Poizat, M., Magnenat-Thalmann, N., Stoddart, A., Thalmann, D., "Mixing Virtual and Real scenes in the site of ancient Pompeii", Computer Animation and Virtual Worlds, 16(1):1-24, John Wiley and Sons Ltd, 2005.

[6] Egges, A., Papagiannakis, G., Magnenat-Thalmann, N., "Presence and Interaction in Mixed Realities", The Visual Computer, Springer-Verlag, 23(5):317-333, 2007.

[7] Ponder, M., Papagiannakis, G., Molet, T., Magnenat-Thalmann, N., Thalmann, D., "VHD++ Framework: Extendible Game Engine with Reusable Components, for VR/AR R&D featuring Advanced Virtual Character Simulation Technologies", Proceedings of Computer Graphics International 2003, pp. 96-104, IEEE Computer Society Press, 2003.

[8] Papagiannakis, G., Magnenat-Thalmann, N., "Mobile Augmented Heritage: Enabling Human Life in ancient Pompeii", International Journal of Architectural Computing, Multi-Science Publishing, 5(2):395-415, 2007.

[9] Chaudhuri, P., Kalra, P., Banerjee, S., "View-Dependent Character Animation", Springer-Verlag, London, ISBN:978-1-84628-591-2, 2007.

[10] Mccarthy, J. M., "Introduction to theoretical kinematics", MIT Press, Cambridge, MA, USA, ISBN:0-262-13252-4, 1990.