Crowd Simulation for Virtual Heritage

*Branislav Ulicny and Daniel Thalmann**

Virtual Reality Lab, EPFL, CH-1015 Lausanne, Switzerland

Abstract

Virtual heritage reconstructions usually focus on displaying architecture or artefacts, with virtual humans playing only minor roles. In the real world, however, such places are, or have been, populated. We introduce a virtual crowd into the real-time 3D reconstruction of a complex heritage edifice to increase the realism of the reconstructed scene. We create a virtual population of worshippers performing morning Namaz prayer inside a virtual mosque. We use a rule behavior system allowing flexible representation of a complex scenario relatively easily adaptable to different edifices or different numbers of persons.

Keywords: crowd simulation, virtual heritage, multi-agent systems, behavioral rules

1 Introduction

With the increase of computers' speed and availability of powerful and affordable graphic cards in the recent years, virtual reality applications became feasible in a broader range of disciplines. Virtual 3D reconstructions of ancient towns, buildings, artefacts, or natural landscapes provide an interesting way of preserving world's cultural and natural heritage.

Virtual heritage reconstructions usually focus on creating visual representations of monuments or buildings, such as cathedrals, with virtual humans playing only a minor role. Usually there is a single "tour guide" [3][4], or virtual humans are more artefacts than living beings such as Xian terra-cotta warriors [5][9] or Egyptian mummies [2]. In the real world, however, most of the reconstructed places are, or have been, populated by smaller or larger number of people - worshippers prayed in cathedrals, druids performed their ceremonies in Stonehenge, gladiators combated in Colosseum in front of spectator crowds.

Although early virtual heritage works have been critised for their lack of visual realism [1], nowadays advances in computer hardware and sophisticated 3D modelling packages allow to create compelling visualisations of static objects. Yet, while photo-realistic architecture reconstructions can be impressive, most of the times they lack dynamic elements such as virtual humans or animals.

We aim to increase the realism of the reconstructed edifices by recreating life inside architectural models. Our work is done in the context of the CAHRISMA project, which aimes to create Hybrid Architectural Heritage, where not only visual, but also

*E-mails: Branislav.Ulicny@epfl.ch, Daniel.Thalmann@epfl.ch

Fig. 1: Crowd performing praying in the Sokullu mosque.

acoustical aspects of the heritage are reconstructed. We integrate a virtual crowd simulation [7] into a real-time photo realistic simulation of complex heritage edifices [6]. We simulate a crowd of virtual humans able to move and interact within a virtual environment. We create a virtual population of worshippers performing morning Namaz prayer inside a virtual mosque (see Figure 1).

2 System design

Our crowd simulation is built as a part of the VHD++ development framework. The VHD++ framework provides components, which allow, for example, loading and displaying of VRML models, animating H-Anim compatible humanoid models, or 3D sound playback. The application is constructed as a particular set of these software components, working on a particular set of data.

The crowd component is responsible for the generation of behaviors of virtual humans. It allows to initialise a population of virtual human agents and then generates in the real-time sequences of actions, such as playing of pre-recorded body or facial animation keyframed sequence, walking to particular location using walking motion model, or playing a sound.

The behavior of the agents is computed by a combination of rules and finite state machines: on the higher level, rules select behaviors appropriate for the state of the simulation; and on the lower level, finite state machines generate sequences of actions. The state of the agent is constitued by an arbitrary number of attributes. Events provide a mean of communication between agents, perception of the environment, and interaction with objects and the user. An important part of the system is a model of the environment containing, for example, semantic information about the space, such as location of doors, or areas accessible for walking and praying. More detailed description of the employed behavior model can be found in [7].

The crowd module allows construction of scenarios defined by a set of behavioral

rules, such as a crowd of worshippers coming to the mosque and performing the religious ceremony.

3 Scenario creation

In the context of the CAHRISMA project, our goal was to enhance the realism of reconstructed Sinan's mosques by adding realistic animated virtual humans to architectural models. The scenario of a morning Namaz prayer was selected as a representation of a typical activity taking place in simulated mosques.

We based the reconstruction of the scenario on video and audio recordings of the actual ceremony. As the first step, we created a structured transcription of the ritual, serving as a further guide for the creation of behavioral rules and the selection of motion sequences and sound clips to be recorded.

VRML models of virtual humans have been created using 3D Studio Max modelling package using a custom plug-in for creating H-Anim compatible hierarchies. As there is a fixed number of polygons possible to be displayed for a particular frame update rate, the polygon count was the constraining factor for creation of human models. The fixed number of polygons have to be divided between the model of the scene and the models of the humans. We opted for creating human models of different complexities ? more complex ones, with around three thousand polygons, for persons with specific roles in the ceremony, such as imam or muezzin, and less complex ones, with around one thousand polygons, for the rest of worshippers. The higher polygon count of "more significant" human models comes mainly due to higher complexity of their facial area, as these models have roles requiring them to perform facial animations. Sounds have been extracted from the audio recording of the ceremony, and corresponding FAP facial animations have been created.

Motion sequences of various parts of the ceremony have been produced by a motion capture of a person performing Namaz prayer. For a convincing scenario, crowd should not look too uniform in the motion: a higher level of variety is needed [8]. Because it would not be feasible to record all motions for every human, we reused a single motion captured animation clip for multiple virtual humans, and the illusion of variety was created by a rule system.

The rule system has two main roles: first, it is responsible for orchestration of the scenario; and second, it helps with creating the variety by generating slightly different commands for the different agents, even while they are executing the same set of rules. Tools for achieving variety are: differences in the durations of the actions; slightly shifted starts of the actions; and the selection of the animations corresponding to the particular action randomly from a set of similar animations.

Most rules are shared by all the agents; some agents with specific roles, such as imam or muezzin have additional, role specific, rules. Sharing a significant number of rules by many agents proved to be important for the development and manageability of the simulation, as eventual changes in the scenario did not have to be propagated to many places in the rule set, leaving less space for errors and speeding-up the development.

Behavioral rules in conjunction with the model of the environment provide a flexible way of representing complex behaviors. One of the requirements of the CAHRISMA project was, that several mosques would be reconstructed. Each mosque, however, is different, with a different size and a different layout of the building. So even while on the higher level of description, the scenario of a morning Namaz is the same for each

Fig. 2: Crowd enters Sokullu mosque

mosque, lower level details, such as the exact location and exact timing of the actions, are different.

To construct these multiple similar scenarios by a linear script would require unnecessary repetitive work prone to errors. The rule system has the advantage over the simpler script to allow specifying behavior on the level of logical units instead of absolute actions with absolute timings. For example, the synchronization of worshippers performing different steps of the praying sequence (as required by the ceremony) is done by inter-agent communication via events. Each person is announcing the end of the current step; a designated leader then observes when everybody finishes the current step, and gives a command to proceed with the next one. Such representation of the behavior is independent of the number of involved persons, or a particular length of the animation clip.

Further flexibility comes from the use of the environmental model: the rules operate on semantic information about the scene rather than on absolute coordinates of the locations. For example, one of the rules states that before starting praying, the agent has to pass through the door of the mosque, and then to arrive to the area designated for praying (see Figure 2). This rule is equally valid for any of the mosques, as the model of the environment supplies the correct absolute coordinates of the door and the praying place.

4 Conclusions and future work

In this paper we presented our work on crowd simulation for virtual heritage. We introduced a virtual crowd into the real-time 3D reconstruction of a complex heritage edifice to increase the realism of the reconstructed scene. We created a virtual population of worshippers performing morning Namaz prayer inside a virtual mosque. We used a rule behavior system allowing flexible representation of a complex scenario relatively easily adaptable to different edifices or different numbers of persons.

We are currently working on extending interaction possibilities of the system using multi-modal interface, where the user can interact with the crowd for example by voice recognition. Another possible extension is the use of the level of details, which would allow to increase the number of visualized virtual humans.

5 Acknowledgements

We are grateful to Mireille Clavien, Rachel Cêtre and Stéphanie Noverraz for the design of the virtual humans and the mosque; Michal Ponder, Bruno Herbelin, George Papagiannakis and Tom Molet for their support with VHD++ framework. This work has been supported by the Swiss National Research Foundation and the Federal Office for Education and Science in the framework of the EU INCO-MED CAHRISMA project.

References

[1] A. C. Addison. Emerging trends in virtual heritage. *IEEE Multimedia*, 7(2):22–25, 2000.

[2] G. Attardi, M. Betrò, M. Forte, R. Gori, A. Guidazzoli, S. Imboden, and F. Mallegni. 3d facial reconstruction and visualization of ancient egyptian mummies using spiral ct data. In *Proc. Siggraph'99*, 1999.

[3] V. J. DeLeon. Vrnd: Notre-dame cathedral: A globally accessible multi-user real-time virtual reconstruction. In *Proc. Virtual Systems and Multimedia 1999*, 1999.

[4] T. Fröhlich, B. Lutz, W. Kresse, and J. Behr. The virtual cathedral of siena. *Computer Graphik topics*, 3, 2001.

[5] N. Magnenat-Thalmann, I. S. Pandzic, and J-C. Moussaly. The making of the terra-cotta xian soldiers. *Digital Creativity*, 8(2):66–73, 1997.

[6] G. Papagiannakis, G. L'Hoste, A. Foni, and N. Magnenat-Thalmann. Real-time photo realistic simulation of complex heritage edifices. In *Proc. Virtual Systems and Multimedia 2001*, pages 218–227, 2001.

[7] B. Ulicny and D. Thalmann. Crowd simulation for interactive virtual environments and vr training systems. In *Proc. Eurographics Workshop on Animation and Simulation*, pages 163–170. Springer-Verlag, 2001.

[8] B. Ulicny and D. Thalmann. Towards interactive real-time crowd behavior simulation. *Computer Graphics Forum*, 2002. (to appear).

[9] J.Y.and Zhong Li Zhang Zheng. Virtual recovery of excavated relics. *IEEE Computer Graphics and Applications*, 19(3):6–11, 1999.