Original article

AI and virtual crowds: Populating the Colosseum

Diego Gutierrez [a], Bernard Frischer [b], Eva Cerezo [a,*], Ana Gomez [a], Francisco Seron [a]

[a] *Advanced Computer Graphics Group (GIGA), Engineering Research Institute of Aragon (I3A), University of Zaragoza, Spain*
[b] *Institute for Advanced Technology, Humanities University of Virginia, VA, USA*

Abstract

Computer technologies and digital recreations have been widely used in the field of Cultural Heritage in the past decade. However, most of the effort has concentrated in accurate data gathering and geometrical representation of buildings and sites. Only very recently, works are starting to go beyond that approach by including digital people. The impressive development of computer graphics techniques and computing power, makes now possible the creation and management of virtual environments where a big number of virtual creatures interact and behave in a smart manner.

In this paper we present a novel use of virtual crowds for Cultural Heritage: we use them to predict behaviors, or to help scholars draw more educated conclusions on unknown matters. We specifically present a case study based on an artificial intelligence crowd simulation which is being used by scholars to study the ergonomics of the Roman Colosseum: it was formerly believed to be an excellent people-mover, but currently that belief is seriously questioned, as potential bottlenecks seem to have been detected.
© 2007 Elsevier Masson SAS. All rights reserved.

Keywords: Virtual reality; Artificial intelligence; Crowds simulation

1. Research aims

Virtual Reality technology is increasingly being employed in many areas of the sciences and humanities. In the humanities, however, its use has traditionally been limited to illustration of buildings and sites through digital reconstruction. While this is a very valid approach, especially useful for education, tourism, and conservation, it does not begin to exhaust all the possibilities of this powerful technology.

On the contrary, Predictive Virtual Reality goes beyond mere visualization. It is a tool that can be used to analyze a problem under different scenarios, test different hypotheses, and result in valid conclusions based on those tests. It allows us to recreate the conditions necessary for the experiment in the form of a computer model and to run accurate simulations that would otherwise be impossible to do.

By combining this approach with Artificial Intelligence algorithms and techniques, one can, for example, populate long-vanished buildings and sites with virtual actors that behave correctly according to certain rules encoded in their virtual brains. Crowds can be simulated this way, overcoming the intrinsically dead nature of computer reconstructions and making it possible for scholars to study the problem of a site in a new way.

Our work differs from previous works using crowds in cultural heritage sites in that we use a bottom-up approach, where the consequences of the actions of the virtual agents are uncertain. It is precisely this uncertainty which allows scholars to test different hypotheses and draw conclusions based on the results of the simulation. More precisely, we study the ergonomics of the Roman Colosseum, and question its reputation as an excellent people-mover.

* Corresponding author. Departamento de Informática e Ingeniería de sistemas, Universidad de Zaragoza, Campus Río Ebro, Edificio Ada Byron, C/Maria de Luna 3, E-50018 Zaragoza, Spain. Fax: +34 976761914.

E-mail addresses: diegog@unizar.es (D. Gutierrez), bernard.frischer@gmail.com (B. Frischer), ecerezo@unizar.es (E. Cerezo), seron@unizar.es (F. Seron).

2. Introduction

The impressive development of computer graphics techniques and computing power has made possible the creation and management of virtual environments where a big number of virtual creatures interact and behave in a smart manner. Although the most well-known applications of crowd simulations are those related to the creation of sophisticated visual effects in the film industry, there are many other ones: creation of training environments for soldiers and policemen, study of building evacuation systems, video games development, studies about herd behaviors, sociological simulations….

Evidently, cultural heritage reconstruction is another exciting field of application. Many works have been published to accurately reproduce the past by using digital technologies (see a good reference of them in Ref. [1]). But most of the initial efforts concentrated only on good data acquisition and geometric representation. More recently, researches began exploring light and its influence on the correct perception of the models [2,3]. In later works, a virtual character is introduced, generally acting as a guide. This is of course interesting, but it cannot be forgotten that most of the reconstructed sites should have been inhabited by a great amount of people: congregations praying in the cathedrals, spectators in the theaters, citizens in the streets and squares,… Some works are now considering these issues. One work [4] introduces a small amount of virtual worshippers in a digital mosque. An impostor-crowd is introduced in the reconstruction of Agora, Greece [5],

whereas a virtual audience in an ancient Roman odeon in Aphrodisias is presented in Ref. [6].

In this paper we present a novel use of artificial intelligence to simulate crowds in virtual environments: to predict behaviors that can help scholars draw more educated conclusions on unknown matters. The case study we present is the analysis of the ergonomics of the Roman Colosseum: it was formerly believed to be an excellent people-mover, but currently that belief is seriously being questioned.

The rest of the paper is organized as follows: Section 3 introduces the Roman Colosseum and those characteristics that are more relevant for the ergonomic study carried out. In Section 4 some issues related to the 3D model of the Colosseum are discussed. Section 5 describes the basics of our crowd simulation system. Results so far are presented in Section 6 and finally, conclusions drawn from this work are explained in Section 7.

3. The Roman Colosseum

The Flavian Amphitheater (conventionally known as the "Colosseum") was built in Rome in the 1970s A.D. by the Emperor Vespasian, who dedicated the partially built complex in 79, the year of his death. The main purpose of the Colosseum was to house the gladiatorial games which had come to be a typical feature of Roman culture in the imperial capital and throughout the Roman world. Other events recorded here include mock naval battles, animal hunts, and the execution of criminals. Figs. 1 and 2 show the Colosseum as it looks today.

Fig. 1. The Colosseum today: exterior view.

Fig. 2. The Colosseum today: interior view.

The Colosseum is the biggest amphitheater ever built, it is said to house between 45,000 and 73,000 spectators. A commonplace in modern scholarship is that it was an excellent people-mover. Fig. 3 shows the different levels of the structure and its original names. The spectators accessed to the grades through 80 doors, arranged along the perimeter. There are five levels of seats. The first level is the *Podio* and was reserved to senators, magistrates and the highest priestly positions. In the center, in correspondence to the minor axis, there are two stalls or boxes: the one in the south was the imperial stall and the one in the north was the one reserved to the magistrates. The second level is the *Maenianum Primun*,
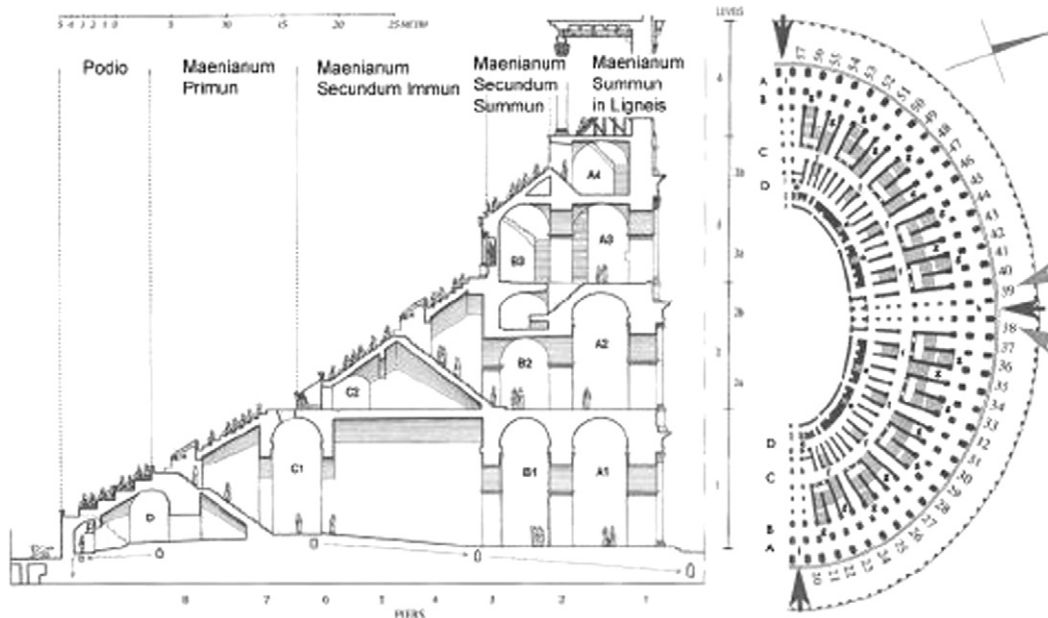

Fig. 3. The different levels of the Colosseum.

reserved to the knights, followed by the *Maenianum Secundum Immun* where citizens and soldiers were seated. The *Maenianum Secundum Summun* was occupied by the rest of the free men and the *Meanianum Summun in Ligneis* was reserved to the plebs, slaves and women.

Not only the seats were distributed according to the social classes of the Roman society, but also the entrances. The exterior arcs of the ground floor facade were all numerated, except the four corresponding to the axes. The emperor's entrance was the one corresponding to the arc not numbered in the south end of the short axis. The magistrates entrance was at the other end of the same axis. Senators accessed the Podio through the adjacent entrances at both sides of the short axis. The rest of the public were spread through the remaining entrances, so that they could reach their seats in the most direct form. There were also two accesses at both ends of the long axis that drove directly to the stage: there were exclusively reserved to the spectacle protagonists. The south east entrance was used to take dead or wounded gladiators out of the Colosseum, and the one in the northwest end was used by the gladiator's parade at the beginning of the games. As it can be seen, the entrance system formed a complex mesh of passages and galleries.

Thanks to the entrance system, social differences were made even more remarkable. According to the standard view, each spectator arrived at the games with a ticket denoting his seat, and even ticket-holders seated in the upper reaches of the *cavea* could supposedly reach their place rather quickly. Egress from the building at the end of the spectacles was also correspondingly quick and efficient. The purpose of the present project is to develop a formal quantitative model to test the validity of this common opinion. The most quantitatively precise version is perhaps that found in Pearson [7]:

"In engineering there are clear affinities between the control of water and of human beings in the mass. In the preliminary designs for the Colosseum, similar foresight was applied to both. One reason why the building has stood for centuries can be attributed to the drainage system hidden beneath the main piers, a carefully constructed line of gullies leading the surplus water from the perimeter to the main sewer. In much the same way the architect devised a system to ensure that his vast amphitheatre would fill and empty perfectly with people. He did this by planning eighty so-called vomitoria – a word which graphically sums up the way the Colosseum spewed out its audience when the show was over-big numbered staircases leading the people to carefully seg-mented rows within the building. These staircases worked so efficiently that it has been calculated that a full audience could leave the building in three minutes flat".

4. Modelling the Colosseum

The first step was to create a suitable 3D model of the Roman Colosseum which could be used to run the simulations. A model had already been successfully developed for previous studies. However, this original model was intended for real-time applications, and some of the important features were missing, such as some stairs, passages, doors and stands' accesses (see Fig. 4). Without an accurate model which included *all* possible passageways and features, the artificial intelligence simulations would be meaningless. In this case, the
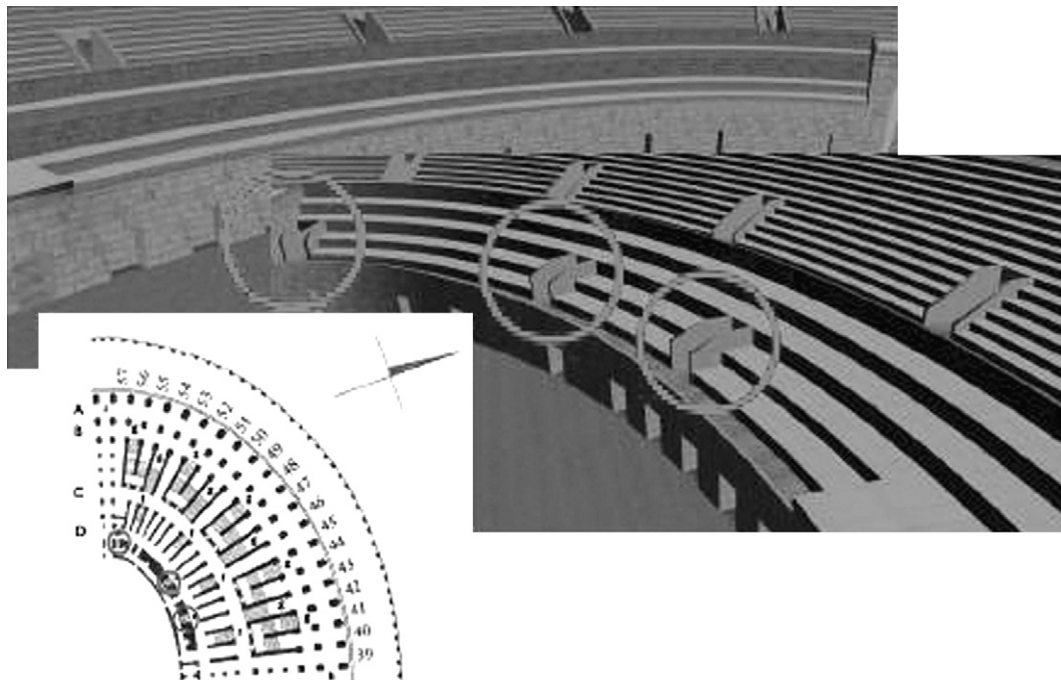

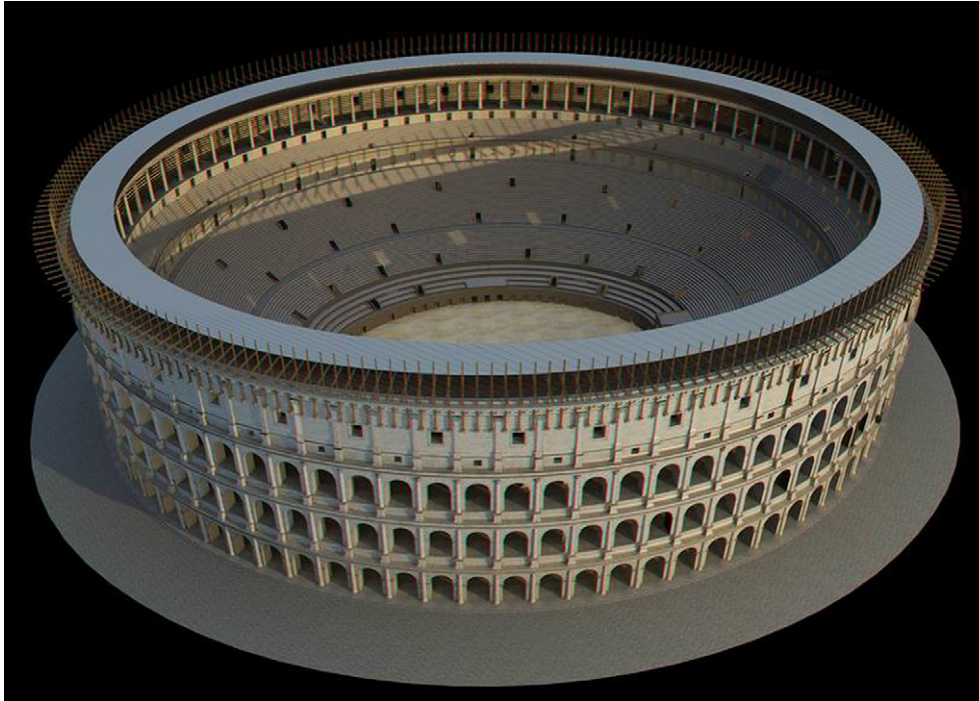Fig. 4. Completing the Colosseum 3D model.

Fig. 5. The digital reconstruction of the Colosseum.

environment was as important as the virtual actors' intelligence. Since the simulation was to be run off-line anyway (computing needs out-ruling real time), all the necessary detail could be added to the model without any restrictions. Fig. 5 shows the final model of the Colosseum.

5. Creating smart crowds

In this section we describe the Artificial Intelligence (AI) framework used in this project, although it is not meant to be exhaustive. A more detailed description on the basics of building an AI engine can be found in Ref. [8]. The approach taken in this work is bottom-up: we build a basic set of rules and study what happens, as opposed to a top-down approach where the goal dictates the behavior rules. The bottom-up approach guarantees that the system is not deterministic, its outcomes cannot be predicted and therefore several unbiased scenarios can be tested. The aim of this work is to develop a multi-agent AI system with scripting capabilities in order to detect possible bottlenecks in the building and to test several hypotheses. The simulation does not need to run in real time; it will be calculated off-line to be then output to a render engine for visualization purposes.

5.1. Virtual agents

In general terms, an agent is a software entity which is placed in an environment and operates under a continuous perception−reasoning−reaction loop. It then first receives as input some stimulus from the environment by using its own perceptual system, it processes it by adding the new information to its previous

knowledge and goals and finally reacts by selecting one in a set of possible actions, which in turn might alter the environment, thus generating new stimuli.

An agent's basic structure is made up of

- Senses: the way it perceives the environment
- Knowledge: a database about itself, its goals and the environment
- Intelligence (behavior): decision-making capabilities based on the knowledge database
- Motor: mechanisms that allow the agent to modify itself and the environment. It represents the agent's capabilities.

An attribute vector for each agent contains information about the agent itself and the environment. This information can be stored, deleted or modified during the simulation, and is the de facto database of the agent. The agents have an adaptive intelligence, where no previous knowledge of the environment is required.

The physical representation of the agent in the virtual world is called avatar. The description of the avatar then includes the software entity known as agent plus its graphical representation (animations, geometry, textures) and its physics (weight, velocity, acceleration). This allows the agent to modify the environment, including another agent.

5.2. Hierarchical Finite State Machines (HFSM)

The Hierarchical Finite State Machines (HFSM) contain the logic of the agent: depending on the state it is in and based on the changes in its attribute vector and/or environment, it

will transition from one state to another, modifying both its attribute vector and the environment if necessary. To do this, the agent has a set of predefined actions, provided by the AI engine (walk, climb the stairs, stop). Even though these actions are predefined, they are generic enough to allow for great flexibility in the behavior of the agents. The term hierarchical simply means that smaller FSMs can be recursively encapsulated as a state of a bigger FSM.

A dynamic event generation system triggers transitions between states. Complex actions can be described by using a scripting language to define them. In a word, the HFSMs should be considered as the brains of the agents.

5.3. Navigation

The virtual environments for the agents are based on 3D Euclidean geometry. A graphics engine handles this layer of the simulation, whereas the AI engine extracts information from the environment and feeds it to the agent (such as there is an obstacle ahead).

For the agents to achieve their goals, three aspects must be considered:

- The sensor system: only the sight has been included in this version, modeled as an angular sector defined by the angle of vision and the visual reach (both parameters can be individually modified for each agent). Other important senses such as hearing are to be added.
- The pathfinding algorithms: Pathfinding (one word) is an AI technique consisting of finding possible routes between two given points. Its implementation is based on the well-known A* algorithm (pronounced A-star).
- Free navigation and obstacle detection: the problem with the pathfinding algorithm is that it computes a route which is not sensible to changes in the environment. To solve this, pathfinding is used along with free navigation

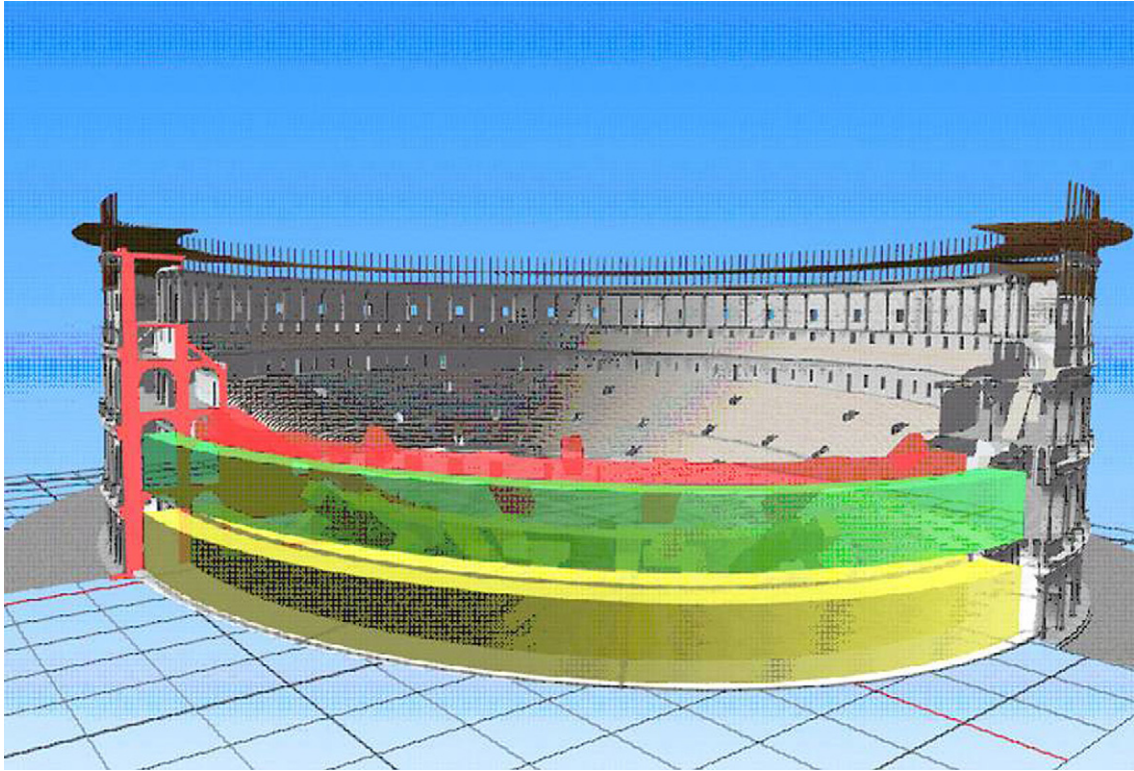Fig. 6. Location of a possible bottleneck in the Colosseum.

Fig. 7. Routes through the Colosseum.

algorithms which allow agents to avoid sudden obstacles returning afterwards to the nearest point in their pathfinding route.

6. The simulation

It is unfortunate that Pearson did not give a source for the calculation. The purpose of the present project is to develop a formal quantitative model to test a novel thesis that states that, for most spectators, passage from the entrance to a seat in the upper levels of the amphitheater and from their seat to the exit was slower than previous scholars lead one to expect. This arises from the detection of some potential bottlenecks in its structure, the most obvious shown in Fig. 6, highlighted by a circle.

Fig. 7 shows a graphic illustrating the circulation routes through the structure. As the illustration makes clear, the routes to the best seats in the lower part of the cavea (yellow and green in Fig. 7), where the citizens of higher status sat, were short, direct and through well-illuminated corridors (see left part of Fig. 8). In contrast, the spectators who had seats at a higher level passed through a relatively low, narrow, and dark corridor (red in Fig. 7). There were no alternative routes: the overwhelming mass of spectators coming to the view the games had, perforce, to pass through this corridor (right part of Fig. 8). Passage through this least spacious and
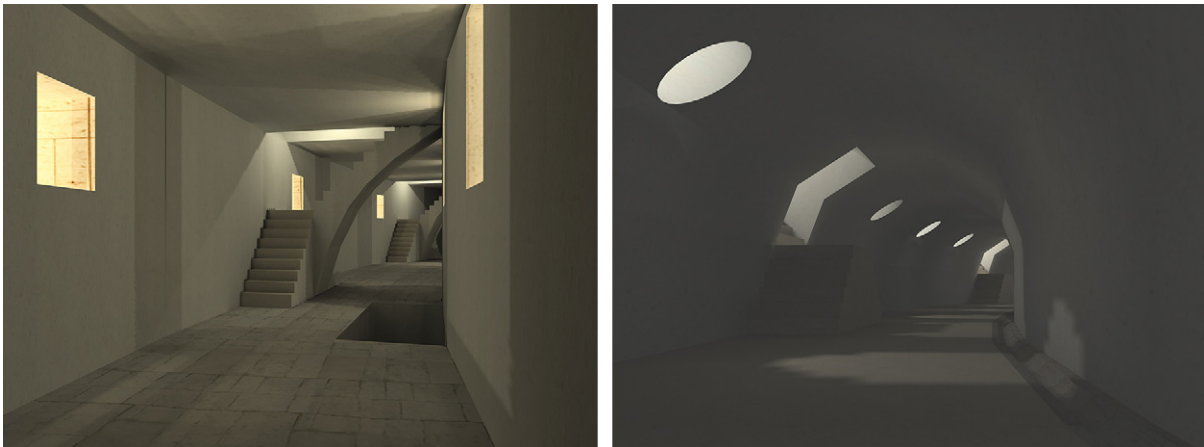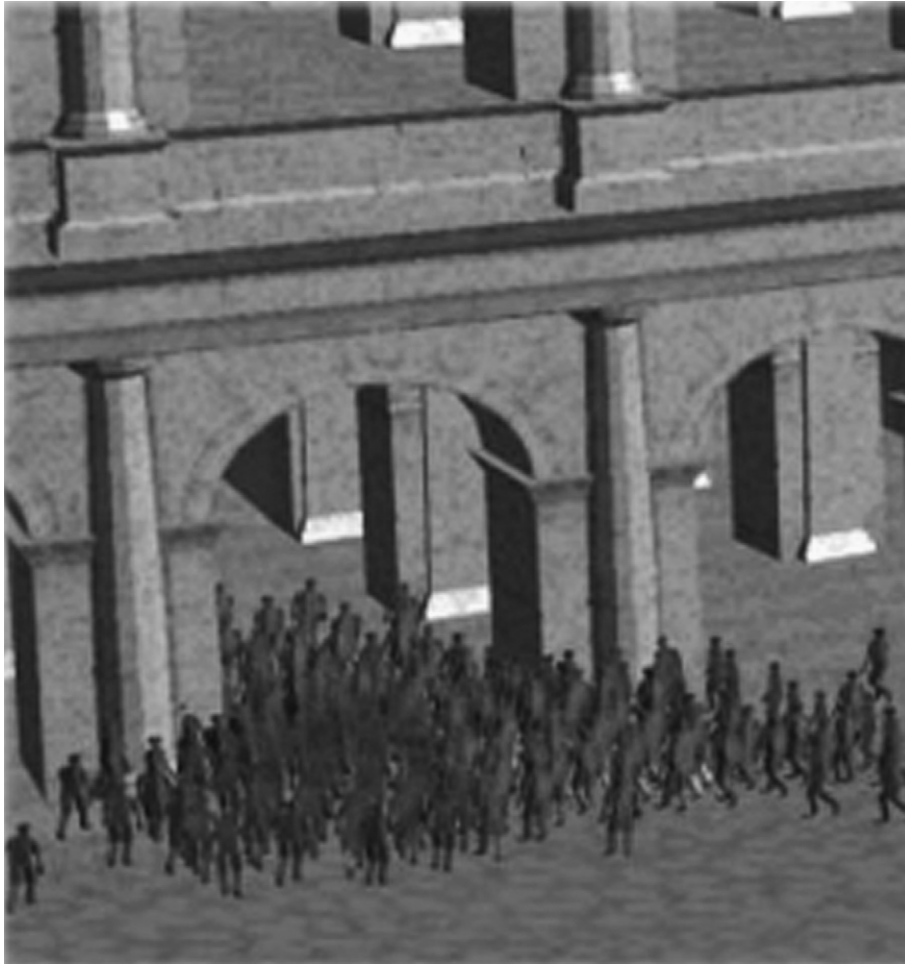

Fig. 8. Corridors inside the Colosseum.

Fig. 9. Digital actors entering the Colosseum.

Fig. 10. Digital actors circulating through the ground floor.

darkest corridor in the superstructure of the Colosseum cannot have been a pleasant experience, no matter the crowd density. One can imagine that it even served to slow down the flow of spectators to their seats (or, at the end of the day's events, to the exits). The present study represents an attempt to take such observations and hypotheses based on eyeballing alone and make them more rigorous and quantitative.

Several simulations have been already tested on a Dual P4 Xeon@2.8Ghz with 2 Gb of RAM. Given that the problem is roughly symmetrical in two axes, only a quarter of the problem has been considered, thus reducing its complexity. Boundary issues between the four quarters of the Colosseum have not been taken into account yet.

A total of 7669 people have been introduced, guided by the AI algorithms. They can react to a dynamic environment (the building plus everybody else, who obviously act as moving obstacles), and know which door they must enter through. Some have some previous knowledge of the Colosseum (as if they had been there before), and some do not. Their goal is simple: to enter by the right door and find their assigned seats. Of these, 251 were considered citizens of higher status, who use the most direct routes to their assigned seats, while the rest were distributed in the different levels of the building. All of them succeeded in finding their way around the building and occupying their place, avoiding obstacles in a dynamic environment where the presence of other agents dynamically changes the environment.

Letting the spectators enter all at the same time at a mean speed of 3 km/h (walking speed), the stand is occupied in about 15/20 min. Some critical points in the circulation system arise, such as the entrances that are crossed by about 400 and 500 people (Fig. 9), or the ground floor gallery, where the knights' path and the path of all the people trying to access higher levels converge (Fig. 10).
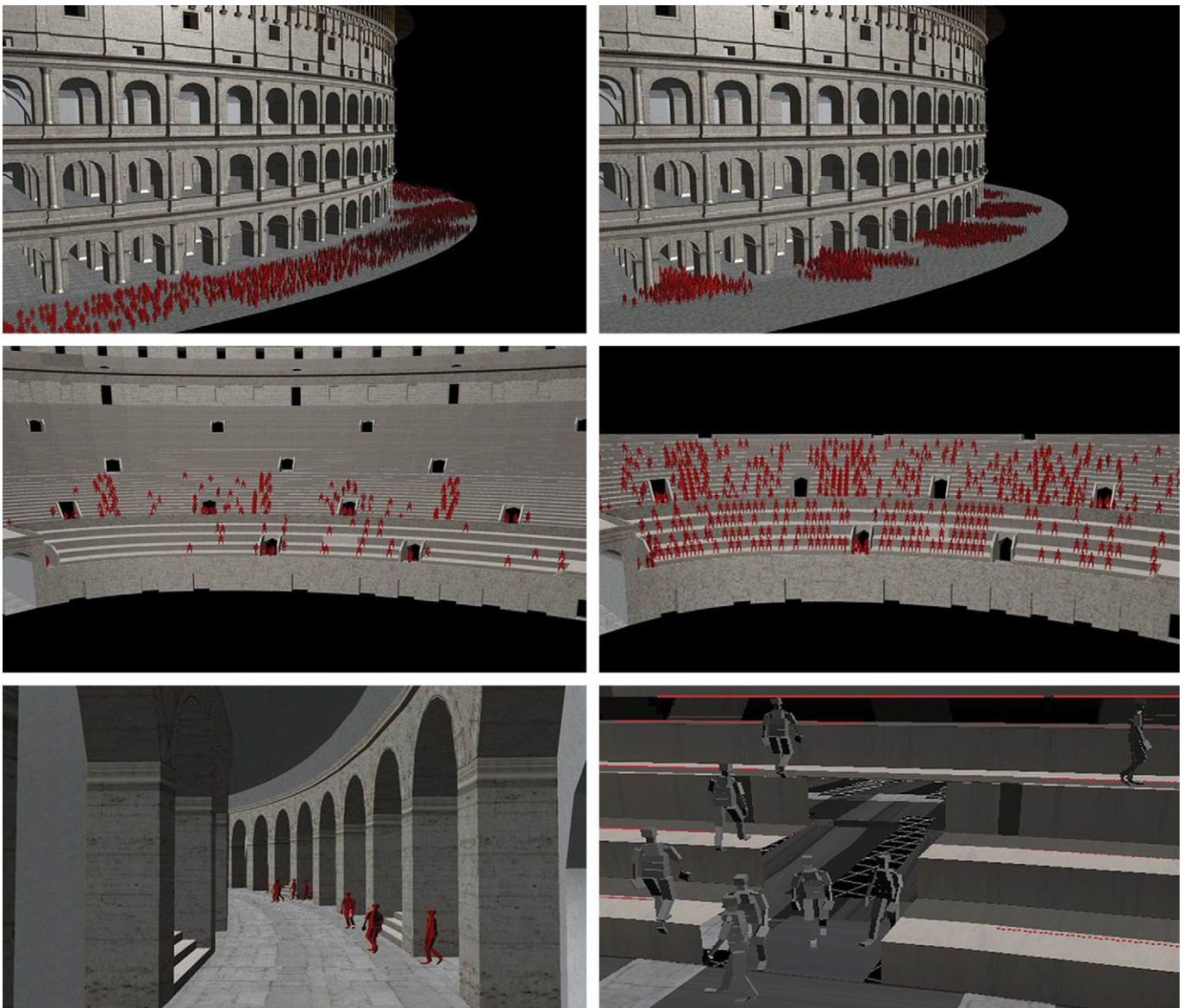

Fig. 11. Some frames of the rendered simulations.

Nevertheless, since scholars do not have a final word on certain key issues, such as how many people tried to enter the Colosseum at the same time, how many doors would remain open for how long or how often the building would be full up to its maximum capacity, many more combinations of hypotheses could be simulated and their outcomes studied. Fig. 11 shows some frames of the rendered simulations.

7. Conclusions and future work

The complexity of the full task of studying the Colosseum is fairly daunting, both because of the sheer size of the Colosseum and the massive amount of agents to be simulated. The simulation is therefore memory-intensive, and advanced optimization strategies must be developed in order to be able to scale the problem to its full dimension. A revised model of the Colosseum was created, based on an original model [9]. This model was adapted to the needs of the simulation: the adaptation process included adding a few missing passages or simplifying the mesh when it was too detailed for the purposes of the project. The behavior of virtual agents has been modeled, by using a continuous perception−action scheme and Hierarchical Finite state Machines. Approximately 8000 synthetic actors, governed by Artificial Intelligence (AI) algorithms, enter the Colosseum through the proper entrance, find their way around, and walk to their pre-assigned seats. The AI is based on state machines, under a perception−reasoning−action scheme. Non-deterministic behaviors can be added to a few random actors, or the characteristics of a given percentage can be altered to observe the effect on the crowd movement.

As predicted, several bottlenecks were detected in the simulation, although the results are not definite yet: different hypotheses need to be tested in order to draw more solid conclusions. Given for instance that the shows lasted the whole day, it is doubtful that all the people would try to enter the building at roughly the same time, and most likely the building would only be at its full capacity during selected fights. Several timings for entering the building will therefore be tested, and conclusions drawn. On the other hand, it is likely that, in whatever order spectators entered the building, the great majority left immediately upon the end of the last event. We will therefore also test the problems that occur when people left their seats and exited the amphitheater. One obvious remaining task is therefore to achieve a complete simulation in the whole building with approximately 50,000 agents. They can also be more varied in size, speed of movement, knowledge of the environment…. Small groups which will tend to advance together can also be added, instead of all individuals. In order to detect bottlenecks more precisely, rendering 3D animations is not really necessary; virtual ''people counters'' will be placed at key spots of the Colosseum instead, and measures of people flux will be visualized in false color maps. This way, it will be easy to identify the suspected bottlenecks just by looking at high-stress areas in the map. Animations can be rendered a posteriori from a selected point of view once the interesting area of conflict is known.

The combination of Virtual Reality technology and Artificial Intelligence algorithms is very promising, and can be used as a new tool for experimental architectural and urban history. We suspect that we have just scratched the surface of its potential.

Acknowledgements

References

[1] J. Barcelo, M. Forte, Donald H. Sanders, Virtual Reality in Archaeology, Archeopress, Oxford, 2000.

[2] I. Roussos, A. Chalmers, High fidelity lighting of knossos, in: IV International Symposium on Virtual Reality, Archaeology and Intelligent Cultural Heritage, VAST (2003), 2003, pp. 47−56.

[3] D. Gutierrez, F. Seron, J. Magallon, E. Sobreviela, P. Latorre, Archaeological and cultural heritage: bringing life to an unearthed muslim suburb in an immersive environment, Journal of Cultural Heritage 5 (1) (2004) 63−74.

[4] G. Papagiannakis, A. Foni, N. Magnenat-Thalmann, Real-time recreated ceremonies in VR-restituted cultural heritage sites, in: Proceedings CIPA'03, 2003.

[5] L. Heigeas, A. Luciani, J. Thollot, N. Castagne, A physically-based particle model of emergent crowd behaviors, in: Proceedings Graphikon'03, 2003.

[6] P. de Heras Ciechomski, B. Ulicny, R. Cetre, D. Thalmann, A case study of a virtual audience in a reconstruction of an ancient roman odeon in aphrodisias, in: V International Symposium on Virtual Reality, Archaeology and Intelligent Cultural Heritage, VAST (2004), 2004, pp. 47−56.

[7] J. Pearson, Arena, The Story of the Colosseum, London, 1973.

[8] J.E. Laird, M. van Lent, Developing an artificial intelligence engine, in: Proceedings of the Game Developers Conference, 1999, pp. 577−588.

[9] B. Frischer. http://cvrlab.org.