2010 Second International Conference on Games and Virtual Worlds for Serious Applications

Levels of Interaction: A User-Guided Experience in Large-Scale Virtual Environments

David Panzoli #1, Christopher Peters #, Ian Dunwell #, Stéphane Sanchez †2, Panagiotis Petridis #, Aristidis Protopsaltis #, Vincent Scesa ‡3 and Sara de Freitas #4

Serious Games Institute
Coventry University Technology Park
Cheetah Road, Coventry, CV1 2TL
West Midlands, UK
[1] DPanzoli@cad.coventry.ac.uk
[4] Sara.deFreitas@coventry.ac.uk

† *IRIT-CNRS, UMR5505*
University of Toulouse
2 rue du Doyen-Gabriel-Marty
31042 Toulouse, France
[2] Stephane.Sanchez@univ-tlse1.fr

‡ *Instinct[Maker]*
Toulouse,
France
[3] v.scesa@instinct-maker.com

***Abstract*—This paper investigates a range of challenges faced in the design of a serious game, teaching history to a player immersed in an 'open' virtual environment. In the context of this paper, such an environment is described as an exploratory, expansive virtual world within which a user may interact in a non-linear, situated fashion with both the environment and virtual characters. The main contribution of this paper consists in the introduction of the levels of interaction (LoI), a novel framework designed to assist in the creation of interactions between the player and characters. The LoI approach also addresses the necessity for balancing computational efficiency with the need to provide believable and interactive virtual characters, by allowing varying degrees of animation, display and, ultimately, interaction detail. This paper demonstrates the challenges faced when implementing such a technique, as well as the potential benefits it brings.**

I. Introduction

This paper considers the challenge faced when seeking to design a serious game populating a large 3D environment, like the high-fidelity model of ancient Rome from the Rome Reborn project[1], where the player is taught history by interacting with autonomous virtual Romans. It builds on previous work at Coventry University [2] that demonstrates the potential of game techniques for cultural heritage experiences, outlining the problems encountered when integrating a substantial number of different state-of-the-art techniques.

(a) (b) (c) (d)

Fig. 1. (a) The Pompeii project aims to populate a virtual replica of the buried city of Pompeii, Italy. Each of the virtual Romans has a different shape, size, or skin colour. (b) The Pennsylvania station is populated with autonomous pedestrians. They form lanes when passing through narrow doors or corridors. (c) The Metropolis project aims to simulate crowds of virtual humans in a city with particular attention to the role of user visual and auditory perception when interacting within the environment. (d) In the game Fable, the interaction abilities are represented by a coloured glowing halo around the NPC.

II. Related work

A significant volume of literature exists around the study of projects aiming to populate a virtual environment with a crowd of characters. Crowds are desirable for a range of purposes, such as bringing life and immersiveness to a historical place, accurately simulating the behaviour of human people, or application-driven scenarios, such as emergency evacuations [3].

One of the challenges of the The Pompeii project [4] is to populate the 3D reconstruction of the city of Pompeii with a realistic crowd of Romans. The project places a particular attention on diversity; of the crowd (figure 1.a). first, as the virtual Romans are modelled using a variety of body, face and clothing models, but also of the environment, as different city areas are differentiated (e.g. poor versus wealthy). Another interesting feature includes the implementation of "places of interest", wherein certain salient city features can attract the attention of passers-by.

The Pennsylvania station project [5] aims to populate the historical reconstruction of the famous New York city railroad station by the same name. Shao and Terzopoulos predominantly focus upon emulating the rich variety in behaviour of the characters (figure 1.b). The characters walk toward individual goals in the station. They can rest, sit on public benches, talk with each other, queue in front of ticket desks, and so on. To further accentuate the realism, live events take place randomly, such as attracting the attention of passing characters.

978-0-7695-3986-7/10 $26.00 © 2010 IEEE
DOI 10.1109/VS-GAMES.2010.27

Other projects consider the role of the human viewer in more detail when synthesising virtual crowds. The Metropolis project [6] investigates visual and auditory perception of crowds in city environments (figure 1.c). For example, at far distances from the camera, it may not be necessary to animate or display models at the same degree of detail as when closer.

Commercial games are worthy of interest as well, since the interactions between the user and the 'non-player' characters (NPCs) have a critical impact on the game-play and the overall experience of the player. Although such games vary in nature and context, the approach used to manage the interactions between the player and the NPCs is often similar. Traditionally, the player is not meant to interact with every character in the game. There is a distinction between characters that have been placed at specific location for interaction and other characters that are simply bringing more life to the area.

Different means are used to signify to the player the characters ability for engaging in more detailed interaction. In *Fable*, from Lionhead studios, halos are used to highlight such characters (figure 1.d). Additionally, when the player enters a range of proximity to the character, a sign appears above them to specify the kind of interaction expected. Although not every NPC can trigger an interaction, all the characters in the crowd play nonetheless a role in the immersion, performing mundane or unrewarding tasks (*World of Warcraft* from Blizzard Entertainment) or actively reacting to the player's actions (*Assassin's Creed*, Ubisoft; *Grand Theft Auto*, Rockstar Games). Characters performing these tasks are central to creating a believable, immersive world. The crowd itself may even be considered as a character on its own, like the crowd of flocking zombies from *Dead Rising*, Capcom.

In both exploratory and narrative-driven serious games, not only the crowd characters must be imparted with the necessary intelligence to behave realistically, but they must also be capable of interacting with the user in a coherent and consistent manner, to maintain flow and immersion and thus effective learning transfer (for evidence of this relationship, see the works of Csikszentmihaly and colleagues [7]). Our proposed LoI approach enables this level of immersion and interactivity, by means of a framework capable of supporting such interactions, in a computationally and resource-efficient form.

III. Levels of Interaction

A fundamental challenge addressed by the framework described within this paper is the effective modelling of interactions between the player and characters in a large environment, wherein a substantial number of entities are likely to interact in a non-linear and emergent fashion. In order to integrate the interactions between the player and the NPCs, we propose a novel framework called the Levels of Interaction (LoI). The same way the Level of Detail [8] and the Level of Simulation [9] techniques computationally optimize the rendering and the animation depending on the proximity of the camera, the LoI conceptually simplify the interactions between the player and the NPCs. Graphically, the LoI can be represented as auras [10] of increasing complexity centred on the player's avatar (figure 2) and based on a simple social space metric.

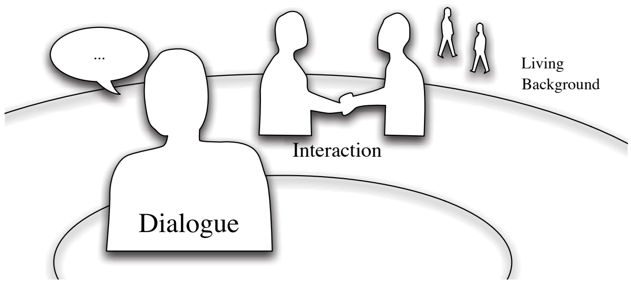

Fig. 2. The levels of interaction (LoI) technique allows different levels of complexity in the interactions between the player and the other characters, from an educational dialogue to a simple living background.

The first level aims to populate the area with an authentic crowd of characters. It provides a *living background* to facilitate the immersion of the player in the environment. Characters evolving in closer surroundings of the player belong to the *interaction* level. These characters pay attention to the player and provide some basic interactions. Finally, a character inside the *dialogue* level interacts with the player in a natural fashion, for example using speech recognition and synthesis.

The LoI framework proposes an elegant and convenient way to represent the types of interactions offered to the player through a single integrated system (figure 3). However, a serious game methodology also entails as primary concern the implications of pedagogy and instructional design paradigms on the model.

Level of Interaction	Living background	Interaction	Dialogue
Distance	Far	Mid-range	Near
LoD	Low	Medium	High
LoS	Low	Medium	High
Granularity	Crowd	Group	Individual
Agent properties	Reactive, Motivational, Opportunistic	Pro-active, Adaptive	Intelligent, Social, Conversational
Example capabilities	Navigation	Complex behaviour, Social attention	Facial expressions, Speech synthesis, Dialogue management

Fig. 3. This table summarises the different features of the characters, for each level of interaction.

A. Living background

The living background layer contributes predominantly to the believability of the environment. At the crowd level, this challenge is addressed within two aspects. Firstly, allowing the characters to adaptively evolve in an informed environment. Secondly, by endowing the characters with motivations and goals to assess the coherence of their behaviour.

An informed or annotated environment contains semantic information, helping the characters to trigger the right behaviour regarding the context. Such information is useful for the general navigation of the crowd [11], defining routes or building entrances, but also for triggering specific behaviours related to specific places, like shops or street happenings. Although their behaviour is guided by these annotations, it is very important – for diversity's sake but also for believability – that the characters maintain their own autonomy. Artificial Life (AL) provides techniques for enhancing their adaptivity. Indeed, increasingly ambitious approaches use AL-driven agents [12], [13] to let some properties emerge instead of scripting them manually: AL-crowds have revealed the ability to model coherent flowing and emergence of lanes, jamming behaviours at constricted points (e.g. doorways) and formation of groups.

To maintain this coherency in time, characters also need to be motivationally-driven. As such, they are guided by internal motivations including internal needs, duties, or a diary, so that everyone of them seems to have a reason to be part of the scene. If the player follows a character, he will witness a citizen living a coherent life. Furthermore, every character has a role and a psycho-sociologic profile [14]. The first defines strong relations between each other – a husband and wife, a master and slave(s), a centurion and soldiers, a merchant and customers. The latter defines soft relations, that mostly depend on the situation – two friends encountering and waving at each other, or stopping their activity to chat; mendicants begging for money as a rich person passes by, considered here as part of the interaction level, as described next.

B. Interactions

Characters inside the interaction level are no longer part of the background since they provide a first degree of interaction with the player. This is revealed by more a complex behaviour, that provides a visual way to teach history, and the ability to dynamically respond to the player.

At the interaction level, LoD and LoS are responsible for rendering the characters more in detail and animating them with smoother motion captures (walking, steering), enhancing their believability. However, teaching history requires the use of even more visually and historically accurate motion captures, providing a first level of educational content by depicting scenes of craftsmanship, art or rites. Figure 4.a shows for example how virtual scenes of acting or dancing, taking place in the Roman theatre of Aspendos, might promote a better understanding of the contextualised social role such places occupied in the ancient societies [15].

(a)

(b)

Fig. 4. (a) Accurate motion captures undoubtedly leverage the pedagogic aspect of the game. (b) The VIBES architecture uses a classifier system to allow the automatic creation of adaptive complex behaviours

Besides, the same technique could be applied to provide accurate depictions of the know-how of a baker, of the meticulous craft-work of a blacksmith, or the social rules involved in a transaction between a merchant and a customer.

The interactive context of the game makes the use of a single lengthy motion-capture inappropriate to represent a complex behaviour. On the contrary, such a behaviour ought to be considered as an adaptive series of actions that the player would be able to interrupt dynamically. Agent technology provides a wide range of controllers for managing the complex behaviour of an animated character like a virtual human. 'Hybrid' controllers[16], [17] are well suited for this kind of task. They combine within a same architecture a database of knowledge representations, for the agent to manipulate objects, and a hierarchical controller able to build complex behaviour from the recursive organisation of lower-complexity behaviours, that can easily be rendered by motion captures (Figure 4.b).

The interaction level also provides the first degree of integration between the player and the crowd, by allowing proactive and non-verbal interactions with the other characters. For instance, passing characters stare at the player, so that he does not feel like a ghost in the crowd. They could also greet the player, if for example the player has met them in the past. More importantly, these interactions provide a way to initiate a dialogue or give the user the incentive to initiate a dialogue. Indeed, non-verbal communication, such as waving or gaze behaviours, can play an important role in the initiation of a verbal interaction [18], for example by attracting the player's attention. The player would then be free to step towards the character, and thus enter the dialogue level.

C. Dialogue

The dialogue level features a close-up dialogic interaction between the player and a character. Since more details of this character can be potentially seen, a highly detailed representation, including for example, facial expressions, gestures, or body stance [19] is recommended. At this level, the role and capabilities of the character are well described by research into Embodied Conversational Agents [20], or ECAs. ECAs are 3D (figure 5.a) or cartoon-like (figure 5.b) characters that attempt to demonstrate many of the same capabilities as humans in face-to-face conversation, including the ability to produce and respond to verbal and non-verbal communication.

(a) (b)

Fig. 5. Depiction of (a) Greta, an Embodied Conversational Agent [20], or ECA, and (b) Lillian the virtual librarian chatbot (©Daden Ltd).

This level particularly supports situated, experiential learning. Therefore, the gaming aspect is largely supported by this level of interaction, and brought to the player via several quests, where each scenario is not explicitly given to him, but built in the course of encounters with a large variety of characters. By talking to different characters, the learner can interact with and reflect upon different aspects of the life and history of the considered period, and for example the differing roles and jobs of individuals within the society.

These dialogic interactions are the very core of this history-teaching serious game, as known pedagogic strategies and techniques can be applied. Namely, the game could benefit from the experience of pedagogical ECAs in order to improve human learning performance. These specific ECAs are virtual instructors that autonomously train a human learner by applying appropriate pedagogical techniques during instruction [21] through scaffolded learning.

IV. Conclusion

In conclusion, this paper has investigated how history could be taught by means of a serious game adopting the LoI approach. The paper contributes towards this long term goal through the provision of a simplified model of the interactions between the player and the virtual characters, in such way as to balance the need for computational efficiency with the need to create believable, immersive environments. The researchers of the SGI have initiated the development of a demonstrator, integrating the Rome Reborn model of the city into the Quest3D game engine. Further development is currently been undertaken in order to implement elements of the LoI framework and investigate low-level and technical implications of it.

As no indication of distance is given in the model, future work intends to study the boundaries of each level of interaction, and the transition between two levels. More specifically, as the different complexity of each level implies a different computation requirement, we think the LoI could provide an additional leverage on the graphical flow of the game.

References

[1] G. Guidi and B. Frischer, "Virtualizing ancient rome: 3d acquisition and modeling of a large plaster-of-paris model of imperial rome," in *Videometrics VIII*, San Jose, California, USA, 2005, pp. 119–133.

[2] E. Anderson, L. McLoughlin, F. Liarokapis, C. Peters, P. Petridis, and S. de Freitas, "Serious games in cultural heritage," in *10th VAST International Symposium on Virtual Reality, Archaeology and Cultural Heritage (VAST'09), VAST-STAR, Short and Project Proceedings*, Malta, September 2009, pp. 29–48.

[3] N. Waldau, P. Gattermann, H. Knoflacher, and M. Schreckenberg, Eds., *Pedestrian and Evacuation Dynamics*. Springer-Verlag, 2007.

[4] J. Mam, S. Haegler, B. Yersin, P. Mller, D. Thalmann, and L. Van Gool, "Populating Ancient Pompeii with Crowds of Virtual Romans," in *Proceedings of the 8th International Symposium on Virtual Reality, Archeology and Cultural Heritage - VAST*, 2007.

[5] W. Shao and D. Terzopoulos, "Autonomous pedestrians," *Graph. Models*, vol. 69, no. 5-6, pp. 246–274, 2007.

[6] R. McDonnell, M. Larkin, S. Dobbyn, S. Collins, and C. O'Sullivan, "Clone attack! perception of crowd variety," in *SIGGRAPH '08: ACM SIGGRAPH 2008 papers*. New York, NY, USA: ACM, 2008, pp. 1–8.

[7] M. Csikszentmihalyi and R. Kubey, "Television and the rest of life: A systematic comparison of subjective experience," *Public Opinion Quarterly*, vol. 45, no. 3, pp. 317–328, 1981.

[8] C. O'Sullivan, J. Cassell, H. Vilhjalmsson, J. Dingliana, S. Dobbyn, B. McNamee, C. Peters, and T. Giang, "Levels of detail for crowds and groups," *Computer Graphics Forum*, vol. 21, no. 4, pp. 733–742, 2002.

[9] J. Pettré, P. d. H. Ciechomski, J. Maïm, B. Yersin, J.-P. Laumond, and D. Thalmann, "Real-time navigating crowds: scalable simulation and rendering: Research articles," *Comput. Animat. Virtual Worlds*, vol. 17, no. 3-4, pp. 445–455, 2006.

[10] C. Greenhalgh and S. Benford, "Massive: a collaborative virtual environment for teleconferencing," *ACM Trans. Comput.-Hum. Interact.*, vol. 2, no. 3, pp. 239–261, 1995.

[11] N. Farenc, R. Boulic, and D. Thalmann, "An informed environment dedicated to the simulation of virtual humans in urban context," in *Eurographics*, 1999, pp. 309–318.

[12] C. W. Reynolds, "Flock, herds ans schools: a distributed behavioural model," in *SIGGRAPH'87 vol. 21(4) of Computer Graphics*. Anaheim (USA): ACM Press, 1987, pp. 25–34.

[13] T. Lightfoot and G. Milne, "Modelling emergent crowd behaviour," in *The Australian Conference on Artificial Life (ACAL)*, Canberra, ACT, Australia, University of New South Wales, 2003, pp. 159–169.

[14] B. MacNamee, S. Dobbyn, P. Cunningham, and C. O'Sullivan, "Men behaving appropriately: Integrating the role passing technique into the aloha system," in *Proceedings of the AISB'02*, 2002, pp. 59–62.

[15] N. Magnenat-Thalmann, A. E. Foni, and N. Cadi-Yazli, "Real-time animation of ancient roman sites," in *GRAPHITE '06: Proceedings of the 4th international conference on Computer graphics and interactive techniques in Australasia and Southeast Asia*. New York, NY, USA: ACM, 2006, pp. 19–30.

[16] T. Conde and D. Thalmann, "An integrated perception for autonomous virtual agents: active and predictive perception: Research articles," *Comput. Animat. Virtual Worlds*, vol. 17, no. 3-4, pp. 457–468, 2006.

[17] S. Sanchez, H. Luga, Y. Duthen, and O. Balet, "Bringing autonomy to virtual characters," in *Fourth IEEE International Symposium and School on Advance Distributed Systems. Published in Lecture Notes in Computer Science, Springer*, vol. 3061, Mexico, 2004.

[18] C. Peters, "Direction of attention perception for conversation initiation in virtual environments," pp. 215–228, 2005.

[19] E. Bevacqua, M. Mancini, and C. Pelachaud, "A listening agent exhibiting variable behaviour," in *IVA '08: Proceedings of the 8th international conference on Intelligent Virtual Agents*, 2008, pp. 262–269.

[20] E. Bevacqua, M. Mancini, R. Niewiadomski, and C. Pelachaud, "An expressive eca showing complex emotions," in *AISB'07 Annual convention, workshop "Language, Speech and Gesture for Expressive Characters"*, Newcastle UK, 2007, pp. 208–216.

[21] J. T. Doswell, "Pedagogical embodied conversational agent," *Advanced Learning Technologies, IEEE International Conference on*, vol. 0, pp. 774–776, 2004.