

Mediating Performance Through Virtual Agents

Gabriella Giannachi¹, Marco Gillies², Nick Kaye¹, and David Swapp³

¹ Centre for Intermedia, Department of Drama,
University of Exeter, New North Rd, EX 4 4LA, UK,
`g.giannachi@exeter.ac.uk` and `n.kaye@exeter.ac.uk`

² Department of Computing, Goldsmiths, University of London,
New Cross, London SE14 6NW, UK
`m.gillies@gold.ac.uk`

³ Department of Computer Science, University College London,
Malet Place, London WC1E 6BT, UK
`d.swapp@cs.ucl.ac.uk`

Abstract. This paper presents the process of creation of virtual agents used in a virtual reality performance. The performance aimed to investigate how drama and performance could inform the creation of virtual agents and also how virtual reality could raise questions for drama and performance. The virtual agents were based on the performance of 2 actors. This paper describes the process of preparing the actors, capturing their performances and transferring them to the virtual agents. A second set of agents was created using non-professional ‘naïve performers’ rather than actors.

1 Introduction

This paper presents work that investigates both how research from drama can inform the creation of virtual agents and how virtual agents can create interesting findings for drama. A virtual reality scenario was created which featured two virtual agents that were based on actors’ performances and explored the use of acting and non-professional, naïve performance for the creation of an appearance of interaction in an immersive VR theatre. We investigated what aspects of the live performance could be transferred to a virtual agent; what technology was needed to recreate the performance, and what effect the nature of the performance had on the final scenario. By comparison an equivalent scenario was created with untrained, naïve performers. This process builds on work using virtual agents for dramatic performance, starting with the work of Bates [1] to more recent work such as the work of Aylett et al.[2] or Mateas and Stern [3]. It also builds on recent performance work using virtual and mixed reality[4].

2 Context and design of the scenario

The scenario was designed with the intention of performing in a type of VR theatre commonly referred to as a CAVE [5]. In contrast with other mediated



Fig. 1. Images from the scenario.

experiences, the CAVE allows the spectator to inhabit the performance via a more complete immersion of the senses than other media – the environment and the virtual agents are seen stereoscopically in a very wide field of view, and the positions of the spectator’s eyes are tracked so that the displayed environment is always perspective correct. This immersion of the visual senses provides good conditions for the elicitation of a sense of ‘presence’ in the person experiencing the performance (henceforth called the participant) – that is, in some sense they will act in and respond to the simulated environment as if it is real. This notion of presence in VR is subject to various interpretations (e.g. [6–8]). An extension of this notion of virtual presence is the idea of social presence [9, 10], in which the participant’s encounters with virtual agents elicit similar social behaviour to that observed in the real world. For such presence to become manifest, the simulated environment should exhibit some degree of believability to the participant.

A key question for designers of VR simulations, and a specific motivation in the design of our scenario, is how to engender believability. To this end, we wanted to investigate the role of interactivity in enhancing believability (and thus the elicitation of presence) by observing how participants behaved when afforded the opportunity to interact with or become involved in the performance.

The scenario involves an exchange between 2 virtual agents, both of which at some stage attempt to interact with the participant. The context of the performance is an experimental study, similar in many ways to the type of experimental investigation that commonly takes place in a CAVE. The scenario was thus designed as a fictitious psychological experiment conducted by the 2 virtual agents, with assistance from a real agent (the experimenter), who undertakes the technical aspects of preparing the participant for the experiment. Images from the scenario are shown in figure 1.

A further design goal was that any interaction among the virtual agents, the participant and the real agent should be as naturalistic as possible, within the confines of a limited range of possible responses (both physical and verbal) on the part of the virtual agents. For example, the start of the “experiment” is delayed several times due to the virtual agents claiming to be “not ready yet”, and then picking fault with the technical setup performed by the real agent. Such activities are designed to allow the participant to establish a sense of an existing relationship among the 3 agents (2 virtual, 1 real) in the performance. In particular, the virtual agents’ criticism of the experimenter at an early stage

is intended to cause the participant to be more willing to engage with the virtual agents (the subtext being that the experimenter is incompetent and should be ignored). The “experiment” then proceeds, albeit with further delays and interruptions, until a final scene in which the virtual agents appear out of character and they ask the participant direct questions about their experience.

Another important aspect of the scenario is a deliberate attempt to blur the boundary between the preamble and the immersive experience as much as possible, thus challenging the participant’s explicit awareness of “being in an experiment”. To this end, the participant’s first encounter with the virtual agents takes place before they enter the laboratory – they can hear the virtual agents talking to the experimenter while they are still outside the lab.

3 Preparing the Performance

Script . As the final context of the scenario is a virtual reality experiment, it was decided to use a fictional experiment as the basis of the script. This was intended to blur conventional distinctions between virtual and real. The script consists of dialogue and interaction between two virtual agents and an experimenter who is played by the real experimenter. The script also involves interaction with the participant. At times the virtual agents directly address the participant and there are moments when a response from the participant is expected. The participant therefore becomes directly involved in the relationship between the agents. This reveals tensions and mistrust between them, and between the virtual agents and the experimenter, whose competency the agents question and consider.

Rehearsal with actors. The actors were prepared using conventional methods for ‘naturalistic’ theatre performance, based broadly on the work of Stanislavski[11]. This enabled the actors to consistently express and reproduce specific emotional qualities and attitudes through their mode of speech, action and interaction with both the other actors and the hypothetical participant. This preparation included the development of an emotional context and backstory in which their behaviour was accounted for and through which each action to be performed in the scenario may be justified. This process involved, firstly, the detailing of fictional contexts that the script itself does not explicitly reveal, but which the actors use to define and understand the motivation behind the moment-to-moment performance the script demands. For example, during rehearsal it was determined that the male character was anxious about his performance and possibility he was considered incompetent by the female character. This was not explicit in the script but was used to inform the emotional context of the actors performance. Through this process and to give expression to this detail, a ‘Physical Score’ was developed consisting of the movements and actions the actors would perform. These actions were always closely related to both the script and the backstory that had been developed. This process involved numerous iterations over a three-week period. At the end of this the actors were able to produce a consistent performance of plausible behaviour which could convey the

sense of a unified and understandable ‘character’. The rehearsal also removed any extraneous, unintended or irrelevant activity, which might detract from this aim.

While the rehearsal process resulted in the actors having a clear conception of the motivation behind each of the actions they performed, it would be incorrect to suppose that the actors have a complete grasp of the ‘character’ and that their conception of that character is conveyed to the participant in the CAVE experiment. In a conventional theatre performance, contemporary theatre and performance theory would suppose that ‘character’ is constructed by members of the audience from their perception of the actor’s performance. This constructed ‘character’ is a function of the spectator’s reading of the performance rather than the actors’ intention. This becomes doubly important when a performance is mediated, for example through virtual agents. Some elements may be lost (for example, facial expression could not be captured) and others may be inadvertently added (for example noise or errors in the motion data). These elements are unlikely to be interpreted as simple errors, rather they will probably be interpreted as non-verbal cues that contribute to the participant’s understanding of the virtual agent’s behaviour and so the participant’s conception of their ‘character’. One aim of this research is to investigate this effect of mediation.

Naïve performers. In order to further investigate the relationship between an actor’s preparation of a performance, its mediation and the participants interpretation, a second set of motion data was captured. This was from two naïve actors with no professional training or experience. These actors were not prepared in the way described above. Instead they were directed to exactly reproduce the actions and timings of the original actors. They were given precise instructions by the director of the actors’ performance, without having seen the original performance themselves. As a result, the motion capture of the naïve actors provided a precise reproduction of the original actors’ performance realized without the supporting structures and techniques intended to ensure fidelity between script, rehearsal and emotional affect. As a consequence, the structure and detail of the scenario and experimental design were fulfilled, but by performers who produced unself-conscious and undirected behaviours – including moments of distraction, inappropriate positions, stances, gazes, as well as tones and intonations – that had no specific design or purpose with regard to dramatic character. The experiment aims to investigate the participant’s interpretation of these actions, in comparison with the more consistent performance by the actors.

4 Capturing and Animating the Performance

The two performances of the virtual agents were captured through the respective performances of the two actors and the two naïve performers. To capture the physical performance, a Vicon (www.vicon.com) motion capture setup using 32 markers per actor recorded and reconstructed their movements in 3D. This affords the capture of the performers’ movements at a coarse level of detail,

but not finer-grained movements such as lip movements, facial expression or hand gestures. Each motion capture session was also video and audio recorded. The recorded video was used as an aid in the motion capture data cleaning, to resolve ambiguities in marker placement. The audio recording was not of sufficient quality to be used in the virtual scenario, but was used as a reference for the studio recording of each audio clip, and for the accurate synchronisation of each clip with the corresponding animation.

In this way, each performer's actions were stripped of many significant details and distilled down to large-scale physical motions, which nevertheless still conveyed much of the original performance. This filtering-out of detail was in some ways akin to the removal, during the actors' rehearsals, of extraneous detail from the performance, with the principle difference being that the loss of detail through motion capture was indifferent to any desired qualities in the resultant performance. Indeed it subsequently became necessary to recreate much of the lost detail. Separately recorded audio samples were carefully synchronised with the resultant animations; the animations themselves, when stitched together to create the complete performance, were precisely positioned, with great care taken to match the performers' postures.

The motion capture took place within a small working volume with approximately 2.5m x 2.2m floor area. This was slightly smaller than the 3m x 3m virtual room in which the scenario takes place, thus it was necessary to spatially offset the recording of some captures. These offset captures could subsequently be repositioned, for example to allow the virtual agents to move within the entire virtual room, and to exit and enter through a doorway. The scenario was broken down into a number of shorter scenes which could be recorded separately and subsequently blended together. Some care had to be taken to ensure that the performers' positioning and posture at the start of each scene was similar to that at the end of the previous scene. The result is a number of different motion capture clips that needed to be combined together to make the final scenario. These captures were arranged into sequences which were played at run time with automatic smooth transitions between them. The script was largely linear so an animation engine structured around sequentially playing clips of motion data worked well. However, it was also important to support interaction of the virtual agents with the participant. An animation engine was therefore developed that supported both sequential, linear actions and more interactive, non-linear ones. This was done with two methods, Cue points and Interruptions.

Cue points. Throughout the scenario, several cue points were identified as locations on the timeline that we might want to instantaneously jump forward to. Typically this would be the case where the participant has answered a question and we want to move forward in the scenario without an unnatural pause. The initial captures contained a reasonably long (about 15 seconds) response phases in which the actor waited for a response from the participant. The response phases were designed to be long enough that the participant could respond comfortably during that time and that the agent would appear to be listening to the

response. However, if the actual response is shorter than the captured response phase the animation should skip forward to the next action to avoid unnatural pauses. During each response phase the actors maintained relatively stable postures to allow naturalistic blending with the subsequent cue point. The animation engine was therefore able to interrupt the response phase anywhere and smoothly blend into the beginning of the next action.

Interruptions and stock responses. Although the transformation of the original performance into the virtual agents' performance filters away some of the salient detail, it also enables the performance to be reconstructed in novel ways that allow pseudo-spontaneous interaction between the participant and the virtual agents. For example, several points were identified as likely occasions that the participant might direct a question to the virtual agents. To allow the virtual agents to respond in some meaningful way, responses to such predictable questions, as well as several stock responses, were performed and processed in the same way as the core performance. All of the responses were captured so that they had similar start and end postures. The responses are animated in a similar way to the cue points, with smooth transitions from the main sequence to the response. However, these main sequences were less constrained than the response phases used for the cue points. For this reason good transition points were found by searching for points that were minimally different from the start and end of the interruptions. The interruptions were then animated by waiting for the next good transition point, smoothly transitioning to the response and then returning to the same point in the original motion.

5 Conclusion and Further Work

This paper has describe the development of a virtual reality dramatic scenario involving a participants interaction with virtual agents. One of our major interests was to investigate the effect of mediating a performance through virtual agents. In particular what effect this has on the participants' understanding of the virtual agents' behaviours and so conception of their 'character'. Actors were carefully prepared to ensure consistent motivation and behaviour, but the resulting performance was necessarily altered by being transferred to virtual agents. The effects of this alteration will be investigated with an experimental study. In this study the performance by the prepared actors will be compared to the performance by the naïve actors. This will help to understand the affect of aspects such as mediation, preparation, consistency and unself-conscious behaviour on the participant's perception of intention, meaning and so character.

Acknowledgements

Research for this paper was conducted in the context of Performing Presence: from the live to the simulated, a four-year collaboration between the University

of Exeter, University College London and Stanford University funded by the Arts and Humanities Research Council (AHRC) UK. Performing Presence is available at: <http://presence.stanford.edu>.

References

1. Bates, J.: Virtual reality, art and entertainment. *Presence* **1**(1) (1992) 133–138
2. Aylett, R., Louchart, S., Dias, J., Paiva, A., Vala, M., Woods, S., Hall, L.E.: Unscripted narrative for affectively driven characters. *IEEE Computer Graphics and Applications* **26**(3) (2006) 42–52
3. Mateas, M., Stern, A.: The interactive drama façade. In Young, R.M., Laird, J.E., eds.: *Proceedings of the First Artificial Intelligence and Interactive Digital Entertainment Conference*, June 1-5, 2005, Marina del Rey, California, USA, AAAI Press (2005) 153–154
4. Kaye, N.: *Multi-Media: video – installation – performance*. Routledge: London & New York (2007)
5. Cruz-Neira, C., Sandin, D.J., DeFanti, T.A.: Surround-screen projection-based virtual reality: the design and implementation of the cave. In: *Proceedings of the 20st Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH 1993, ACM (1993)* 135–142
6. Lombard, M., Ditton, T.: At the heart of it all: The concept of presence. *J. Computer-Mediated Communication* **3**(2) (1997)
7. Zahorik, P., Jenison, R.L.: Presence as being-in-the-world. *Presence* **7**(1) (1998) 78–89
8. Sanchez-Vives, M., Slater, M.: From presence to consciousness through virtual reality. *Nature Reviews Neuroscience* **6**(4) (2005) 332–339
9. Biocca, F.: The cyborg’s dilemma: Progressive embodiment in virtual environments. *J. Computer-Mediated Communication* **3**(2) (1997)
10. Garau, M., Slater, M., Pertaub, D.P., Razzaque, S.: The responses of people to virtual humans in an immersive virtual environment. *Presence* **14**(1) (2005) 104–116
11. Stanislavski, K., Benedetti, J.: *An Actor’s Work: A Student’s Diary*. Routledge (1938)