ABSTRACT

With advances in virtual reality technology and its multiple applications, the need for believable, immersive virtual environments is increasing. Even though current computer graphics methods allow us to develop highly realistic virtual worlds, the main element failing to enhance presence is autonomous groups of human inhabitants. A great number of crowd simulation techniques have emerged in the last decade, but critical details in the crowd’s movements and appearance do not meet the standards necessary to convince VR participants that they are present in a real crowd. In this paper, we review recent advances in the creation of immersive virtual crowds and discuss areas that require further work to turn these simulations into more fully immersive and believable experiences.

1 INTRODUCTION

Immersion describes the extent to which computer displays are capable of delivering an inclusive, extensive, surrounding and vivid illusion of reality to the senses of a human participant [27]. Increased immersion can result in increased presence in the virtual world. Slater and Wilbur define presence as a state of consciousness, the (psychological) sense of being in the virtual environment. Studies show that increased presence in virtual worlds can yield better training transfer from these graphical environments to real world scenarios [27].

Virtual crowds have potential to enhance Virtual Reality (VR) experiences. The way we perceive a virtual crowd strongly depends on the medium and on our role. Our experience changes based on whether we are a passive viewer (e.g. movies, even with 3D technology) or an active participant (e.g. video games). In VR the goal becomes convincing the participant that he/she is experiencing something real (i.e. being more present in the virtual environment than the real one) and is surrounded by a virtual crowd. Unfortunately this is where failures begin. Participants’ expectation of the virtual crowds are not met and their sense of presence is diminished. The problem is that we as humans observe each other every day and become accustomed to normal behavior and attune to recognizing even small deviations from these norms. Any unnatural artifacts in our crowd simulations become magnified by the first person perspective and immersive technologies of VR.

Since we first explored using presence to evaluate crowd simulations [21], a lot of technologies have been developed. VR hardware has advanced and become readily available as consumer level products. Crowd simulations have also been greatly improved, but are not yet to the level they need to be to be effective in VR. Research focuses on: making crowds appear more purposeful, making characters follow reasonable trajectories, making animations smoother and more natural, improving rendering and so on. Yet when immersed in VR, crowd simulations are still found lacking. Just as we fix one distracting artifact, another appears and rolls us right back into the uncanny valley. We seem to ride an uncanny valley roller coaster. Every time we manage to improve one feature, an issue with another becomes apparent and drastically reduces realism yet again.

So how far are we from achieving our goal of effective crowd simulations in VR? That is the focus of this paper. We will evaluate current crowd simulation research, we will discuss the progress of VR technology over the past eight years, and we will present our thoughts on where additional efforts are needed to bring us closer to our goal.

2 CROWD SIMULATION METHODS

There has been a large number of methods to simulate crowds of virtual characters. Most of the crowd simulation work focuses on simulating root trajectories using a bounding volume of the entire character for collision avoidance and repulsion forces. For a good survey on crowd simulation methods we refer the reader to recent books in this field [13, 19, 30]. Simulating root trajectories provides good results in terms of local movement, but artifacts appear easily in the animations and feet-ground contact, specially as the densities increase and careful feet position and animation blending becomes necessary. In those situations it is easy to spot foot sliding, abrupt changes in velocity and torso orientation, as well as jumps between animations.

These approaches often deal with a small database of animations and either limit the movement of the character to the available animations [15], or include some techniques to interpolate between animations and apply time warping to satisfy constraints between the feet and the environment [20].

Some recent work attempts to achieve more natural results by planning in the foot domain [31, 25] and then computing the right sequence of animations and blending to follow the resulting trajectory [8].

Another aspect of virtual crowds that requires attention is rendering. Often when large crowds are required, it is necessary to use techniques to speed up rendering while minimizing loss of visual quality. But is realistic rendering really required for presence? There have been studies on visual realism affecting presence from the point of view of the environment under different illumination models [32]. From those studies it appears that what matters is the dynamics of the illumination model, i.e. changing shadows and reflections. The quality of the illumination does not seem to affect presence, however it results in greater plausibility and thus participants are more likely to respond to virtual events as if they were real.

This opens the door to evaluating the visual appearance of the crowd. Work by Jarabo et al. [12] studied how crowd movement characteristics and different illumination models affected the perception of crowds. The purpose of this work was to determine how much the illumination models could be simplified for crowd rendering without being noticeable. This could also have an application to rendering crowds in IVEs (Immersive Virtual Environments). Evaluating presence under a variety of crowd rendering techniques [7] would help us to determine the most relevant conditions to use each type of rendering and thus overall save resources while still providing realistic visualization of many characters. For example, in Figure 1 we compare simple cartoonish looking characters used in an immersive virtual crowd experiment against a crowd rendered with a more human-like appearance and shadows that change dynamically as the characters move. The next level of visual appear-
ance would be photo-realistic characters, but this would most likely present problems with the Uncanny Valley.

Figure 1: Crowds rendered with cartoon like characters and no shadows [22] (left) or a more human-like appearance and shadows [6] (right).

3 VR Technology

When considering the potential impacts of crowd simulations on virtual reality and vice versa, we must include both software and hardware. Focusing first on software, the more open release of games engines like Unity 3D [1] and the Unreal Engine [2] have provided crowd simulation researchers access to technologies and methods that previously required great effort on their part. Even if the focus of their research was on steering or stepping behaviors, they needed to develop software for navigation, rendering, etc. in order to show how their methods operate in a full crowd simulation. While not yet ideal and not yet entirely simplistic to utilize, these game engines provide a framework that includes at least rudimentary (and in some elements more advanced) components on which researchers can examine and develop their simulations. Functionality includes, level design tools, navigation meshes, animation playback and blending, inverse kinematics, advanced rendering techniques, and physics engines. Pushes toward open source software has also lessened the initial investment of time and effort previously needed when beginning crowd simulation research.

Furthermore, these engines provide easy access to varied input and output devices. Researchers no longer need to be concerned with developing device drivers and connective software for game controllers and Head Mounted Displays (HMDs). These engines provide automatic connections with consumer level hardware like wireless controllers, trackers (e.g. the Microsoft Kinect) and the Oculus Rift. This promotes much greater experimentation at lower investment costs.

Focusing on VR hardware, it now works and is even common place [26]. CAVEs now have high resolution displays with seamless image alignments. There are even a number of portable CAVEs. HMDs also benefit from high resolution as well as wider fields of view, lighter weight, and lower cost points. There are now simple mounts for smart phones to become HMDs. Head trackers and whole body trackers, like the Microsoft Kinect, are common and for many applications have the required accuracy and ease of use. The availability of these devices makes it much more feasible to place simulated crowds in VR and potentially provide new lines of research in participant interaction with crowds.

Devices such as eye trackers and biometric sensors (e.g. EEGs and SCRs) have also become much more pervasive and are used to evaluate VR applications including the levels of presence of the participants [28]. Using these devices to aid in the determination of a subject’s level of presence in VR worlds with varying crowd models could be a viable avenue for crowd simulation evaluation.

4 State of the Art in Immersive Virtual Crowds

Immersive virtual crowds could offer a great potential to evaluate human behavior in difficult or dangerous situations. Virtual reality offers the possibility of replicating experiments without exposing the participant to a real dangerous situation. For example, there has been work studying human behaviors under a fire emergency in a public place [28]. This particular study was carried out with a few virtual agents with scripted behaviors and animations. Virtual crowds in immersive environments have been proposed as both a validation technique for crowd simulation from an egocentric point of view, and also as a setup to learn about human decision-making in controlled situation [21]. The learned data could be used later on to enhance crowd simulation models.

In order to be able to both validate and extract relevant data, it is essential to create an environment where people feel presence, and thus react in a realistic way to virtual events. In [21] the authors present a study based on presence literature and features that appear often in crowd simulation models. They discuss how those features could enhance or break presence. In a later experiment [22] using a head mounted display, they showed that having some level of interaction with the crowds could enhance the feeling of being there. For example, the virtual agents were able to see the participant to perform collision avoidance, they could also be pushed by the participant in case of high densities. The participant was immersed in a virtual cocktail party where the crowd would walk around and stop at tables with food and drinks. Agents would autonomously decide to stop around tables, turn to face other agents, and play idle behaviors that included hand gestures imitating being engaged in conversation. All of these elements combined with the background noise of plates and conversation helped create a strong sense of presence despite the cartoonist appearance of the characters.

CAVE systems have also been used to immerse a participant in virtual crowd for the purpose of evaluating new crowd simulation methods [3]. In this work the goal was to compare different crowd simulation methods. The participant was standing in a location waiting for a person and could observe the crowds from an egocentric point of view to evaluate their collision avoidance methods. Questionnaires were used afterward to gather information about the participants opinion regarding life-likeness of the crowd, spatial awareness and comfort in the CAVE.

Both the work by Pelechano et al. [22] and Ahn et al. [3] found that breaks in presence or low comfort appear when there is no collision avoidance (i.e. characters moving through the participant) and when the method produces shaking artifacts.

The work by Kyriakou et al. [14] demonstrated that showing collision avoidance between the user and the virtual agents with some basic interaction between them could increase the sense of presence. This kind of interaction also helped to make the virtual environment appear more realistic and lifelike. In their studies they also observed that including interaction with the virtual crowd could enhance presence. The types of interaction used for their work were limited to having characters waving or verbally saluting the participant.

The work by Bruneau et al. [9] studied the participants’ behavior when walking against groups of virtual agents. The goal was to determine what elements of the crowd get people to decide to walk through or around a group of people. The elements of the group that they tested where based on relative motion, size, density or appearance.

Proxemics have also been evaluated with small groups of virtual characters walking towards the participant [17]. Those experiments found that physiological arousal would increase as the characters would approach the participant and also as the number of virtual humans would increase from 1 to 4.

These last two papers had some interesting, if not surprising, results which were that in certain situations the appearance of the virtual humans did not seem to have much effect on the participant’s response. For example the first one [9] was evaluating the participants’ decisions of walking through or around a group of virtual
humans. It turned out that whether the virtual humans were rendered as ordinary people or zombies had no significant differences in the decisions made by participants. And similarly, in [17] they found no significant differences in terms of physiological arousal when being approached by virtual people or cylinders. This opens questions regarding the extend to which realistic visual appearance is necessary and why.

IVE provides a platform to study crowd behavior under controlled conditions [18]. In this paper they discuss and evaluate the use of VR to carry out controlled experiments to obtain data about human behavior in terms of navigation. The long term goal would be to be able to use VR as an aid in the process of developing more accurate microscopic models. The paper raises interesting questions such as: are trajectories in VR similar to those in real live? or, is there any bias induced by the use of locomotion interfaces?

The answers to these questions are still unresolved. For instance, part of the work in this field is focusing on enhancing presence in an attempt to have the participants behaving as in real life. But how close are real responses to virtual events from having participants performing realistic local movement? Given that movement itself is constrained in most VR applications, (i.e. cables in HMD or size of the CAVEs) can we fully trust participant movement? In other words, is presence enough to turn VR into an accurate platform to capture human locomotion? There has been a large amount of work trying to find correlations between human walking in static virtual environments and in real setups that exactly match the dimensions of the virtual one. Lately some researchers have been moving these studies to dynamic environments populated with virtual humans to find the similarities between the real and the virtual world [24].

Despite recent advances in VR and the availability of engines such as Unreal [2] and Unity [1] that facilitate the task of creating virtual environments, there is still a large amount of manual work required to populate those environments in a believable manner. The paper by Spanlang et al. [29] provides a good overview of the difficulties in creating a virtual experiment dealing with embodiment. Embodiment requires the replacement of the participant’s body with a virtual character seen from a first person perspective. To be believable, the body representation needs to follow the movements of the participant. This implies having at least one virtual human, with real-time motion capture, character animation, and realistic rendering. Having a few believable virtual characters already requires a large amount of work, but the problems escalate as the number of characters needed increases, since it is necessary to deal not only with the simulation of each character but also with the interactions between them.

Most of the time these virtual characters follow strictly scripted behaviors and their apparent interaction simply follows the Wizard of Oz metaphor. Animations are motion captured and carefully cleaned and adjusted to the final scenario. In such experiments, there is little room for autonomous behavior, and the virtual humans are like the actors in a movie with little room for interaction with the participant. When we deal with immersive crowds, the job of editing scenarios with such level of detail becomes unfeasible, and thus more research is needed to move presence experimentation into the crowd simulation field.

5 SO WHAT’S MISSING?

In [26] Slater discusses grand challenges in virtual environments, areas that need further research. One topic highlighted is tracking and the need for better body tracking over wide areas with high accuracy and low latency. Such advancements in tracking could positively impact crowd simulation research on two fronts. One, it would enable participants immersed in a virtual crowd to move more naturally and more freely. It could also enable better data collection for input to crowd simulations and to evaluate them.

Slater also discusses plausibility illusion. Participants carry expectations about their environment and the people in it from the real world into the virtual worlds. This calls for more realistic human behaviors including the behaviors of crowds and their interactions with the environment. Said interactions require a high level of domain knowledge, the creation of which is tedious and time consuming. Recent research has explored automating the creation of elements of domain knowledge [23, 5], but additional effort is needed.

The plausibility illusion also calls for researchers to avoid the Uncanny Valley. When participant expectations are not met, presence can be broken. When participant expectations are set too high by one aspect of the simulation, lower quality in another aspect can break presence. Along these lines, there are areas of active research that need continued improvement for crowd simulations in VR to be successful. We need better animation and blending tools, better rendering for large crowds, better stepping and navigation algorithms (particularly in cluttered environments). In fact animation bugs and glitches tend to be very noticeable in VR and very distracting. We also need to calibrate the speed of characters when placed in VR. We have observed that they tend to be perceived as moving more quickly than they should.

According to Slater, another challenge for virtual environments is haptics. While there are an array of haptic devices available, none can replicate all, or perhaps even any, of the touches and feels we experience daily. Slater points out that research in the area of haptics has to contend with and is limited by the laws of physics, making it particularly challenging. A lack of realistic haptic feedback is problematic when immersing participants in crowds. The lack of tactile feedback from even relatively minor collisions with individuals in the crowd may greatly decrease presence. Absence of the forces and pressures normally experienced in dense crowds would greatly decrease a participant’s sense of presence. Research on how to replicate the feeling of being immersed in a crowd is still needed. Do advances in neuroscience hold the solutions?

Another challenge is enabling people naïve to programming to create and modify environments. Current software frameworks like the game engines discussed in Section 3 have begun addressing this challenge. They provide tools for importing models, designing levels, blending animations, designing simple AI, etc. They do not, however, provide tools for the creation and control of virtual human populations. While we have explored authoring the behaviors of populations from a limited number of easy to author action types (See figure 2) [19, 16, 4], additional research is needed.

Figure 2: Simple GUIs for authoring reactive, stochastic, and need-based actions (left) and Microsoft Outlook calendars for authoring scheduled actions (right).

Cummings et al., who studied immersion and presence through a meta-analysis of previous research [11], discussed different types of presence. They indicate that social presence would be increased through better communication and social actors. Certainly this is an element of crowd simulations that could use additional research. Characters in crowds tend to be expressionless and largely non-active. While they may respond to avoid collisions and to a few scripted events, they tend to appear rather lifeless. They possess
limited communication models and few social features. Even the addition of models of attention would provide more lifelike qualities. Characters should look where they are going and at interesting objects and events in the world. Individuals in crowds that meet the eye of the participant from time to time could help to increase social presence. Characters should react to each other and to the participant in a natural manner. Figure 3 shows that even in a rather primitive crowd simulation and VR setup, participants can be positively influenced by even unintentional character actions. To increase social presence, characters should exhibit surprise, curiosity, and even anger. They should exhibit different personalities and attitudes. They should appear...human.

In terms of spatial presence they found tracking level, stereoscopy, and field-of-view to be more impactful than image quality, resolution, and sound. Tracking provides participants with a sense of self-location, better navigation, and additional action possibilities. Stereoscopy provides greater spatial presence through increased depth perception and certainly wider fields-of-view aid in self location. They found that the ability to navigate around the environment is key to increasing spatial presence. Navigation is more difficult in crowds, especially dense crowds. Will including crowds in VR inherently decrease levels of spatial presences or will participants have lower expectations for navigation in crowds and have their levels of presence remain high?

Another element of crowd simulations that has not received much attention is sound. Adding realistic, localized sound to the characters in a simulation may improve the participant’s level of presence. They would be less startled by characters suddenly appearing in their field-of-view. The characters themselves could react to noises in the environment, making them appear more lifelike.

Finally, elements of story and game play within crowd simulations may increase immersion and presence. If participants are enthralled by a story or activity within the virtual environment, they may notice fewer anomalies and be less impacted by them. Perhaps elements of story and game play need to be addressed in crowd simulation research.

6 Conclusion

VR can benefit from advances in crowd simulation techniques and crowd simulation research can benefit from advances in VR technology. Populations of virtual humans can make immersive virtual worlds more realistic, interesting, and compelling. The crowd simulation community continues to explore ways of evaluating and validating their techniques. VR is an ever more viable venue for immersing participants in a crowd and examining their behavioral reactions and overall levels of presence. However, as outlined in this paper crowd simulations must still improve to meet the standards needed to achieve high levels of presence.

Since VR presents an ideal platform to study human behavior in crowds, it is thus important to focus our research efforts in determining how we can achieve realistic immersive crowds. One way that we found very valuable to make progress is to obtain more information from the participants after the experiments. So for instance, once all the results from questionnaires and physiological response have been gathered, we can ask the participants to give us general comments about the experiment. What was the most compelling? What did they find was missing? What broke the feeling of interacting with a real crowd? Their answers can help the researchers to think outside the box by noticing what the average person may be experiencing as opposed to what the researcher feels is missing. It also provides relevant information regarding what the average person may not care about that the researchers do. For example, when playing video games, people get so immersed and excited about the game that they voluntarily decide to ignore the artifacts (i.e. flow [10]). So the player does not stop chasing the enemy just because he/she observed a bad blending, an abrupt change in direction, or some foot sliding. Could participants in a presence environment with an immersive crowd also adapt to the circumstances and decide to ignore certain artifacts? Maybe as long as the virtual humans appear real by talking to them, waving or showing emotions, the participant would treat them as real people.

More research is needed.

7 Acknowledgments

This work has been partially funded by the Spanish Ministry of Science and Innovation and FEDER under grant TIN2014-52211-C2-1-R.

References


Figure 3: A participant in a VR experiment on presence reacting to a perceived wave from a virtual character by starting to wave back (movement of left hand).


