Plausible Methods For Populating Virtual Scenes

*Cathy Ennis, Anton Gerdelan and Carol O'Sullivan
Graphics Vision and Visualisation Group, Trinity College Dublin

Abstract

We describe a series of experiments investigated the perception of urban crowds and the implementation of these results in a real-time crowd system. Our first series of perceptual experiments examined the role of context in locations containing pedestrian crowds, where properties of the location affect the formation of the individuals making up the crowd, i.e. their position and orientation. Our second set of experiments focussed on small groups of individuals. We considered group size and number of dynamic groups, and looked at body motions and audio desynchronisation of individuals in conversational groups. We also discuss the implementation of results from these experiments in a real-time crowd system.
Keywords: Perception, Virtual Crowds

1 Introduction

Virtual crowds are used in many applications across different industries. In the entertainment industry they are used for many purposes; in movies, for example, computer generated crowds are used to ameliorate the costs involved in hiring a large number of extras. Generic crowds are used in video games, particularly in urban environment games such as Grand Theft Auto or The Sims. With increasing demand for realism in video games, however, the need for increased realism in the behaviour of individuals and the crowd as a whole is becoming more important. Humans are inherently good at recognising behaviour patterns in other humans. This applies to both navigational and social behaviours. However, less is known about how we perceive virtual crowds and what is expected of them to create a plausible virtual crowd.

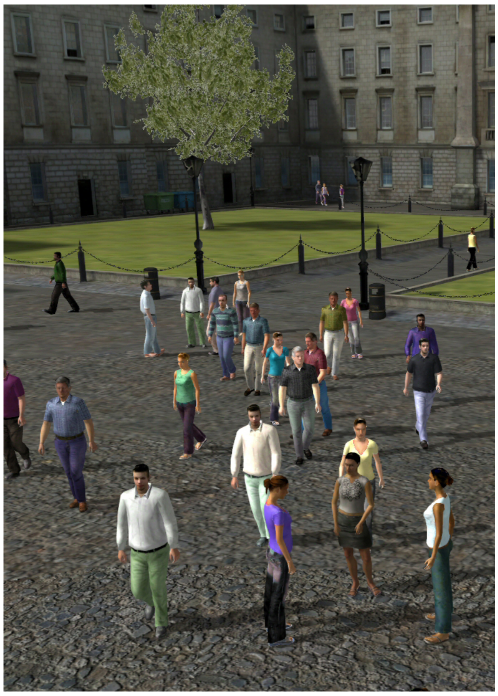

Figure 1: Screenshot of our real-time crowd system showing conversational agents and dynamic agents.

In this paper, we give an overview of several perceptual studies that have been carried out to evaluate methods of creating virtual environments with pedestrian crowds. We also describe how the results from these experiments have been implemented into our real-time crowd system of urban pedestrians with different behaviour roles. One aspect of this research involves populating these environments, with agents in realistic positions and orientations [12, 13]. Another aspect involves main-

*e-mail:ennisca@cs.tcd.ie

taining a sense of realism during simulation of a crowd system, through the addition of plausible groups [11, 14, 15].

2 Background

Previous research in the field of urban planning and civil engineering has highlighted everyday behaviours of pedestrians such as: planning itineraries based on their knowledge and perception of the environment [1]; use of space and walking speed [2]; bi-directional lane formation [3]; collision avoidance [4] and side-stepping [5]. Other researchers have investigated more social behaviours, such as the formation of small conversational groups [6]. Many behaviour models use data collected from real humans to control the actions of virtual agents. Lerner et al. [7, 8] use human motion data collected from video to create a database of example trajectories, one of which is chosen at runtime for a virtual agent. Other models have used human data to create models of gesturing styles of people [9]. Cassell et al. [10] developed a model for conversational behaviour using semantics from text to identify appropriate speech, gestures and facial expressions.

While much research has been carried out regarding into the behaviour of virtual agents, little is known about how we perceive these behavioural models, or the virtual agents themselves. Our previous work has investigated human perception of virtual crowd formations and sensitivity to temporal desynchronisation of conversational gestures between individuals in a group. This will be discussed in further detail in Section 3.

3 Perceptual Experiments

3.1 Static Crowd Formations

Since studies focusing on how viewers perceive and evaluate the plausibility of synthesised behaviours remain scant, we developed an evaluation methodology to allow us to better elucidate this complex issue [12, 13]. Our methodology consists of experiments to evaluate participants' general impressions of static scenes, considering the orientation and positioning of pedestrians. Rather than considering pedestrian properties in isolation, a central theme in our methodology is to also consider the context in which they appear. The term *global context* refers to the type of location in which the pedestrians are situated. Local context refers to properties of different locations, which infer certain behaviours:

- Individuals tend to stay in-bounds of walkable areas
- At any given time, there will be a number of pedestrians walking in groups of two or more, and will be facing the same direction
- Individuals tend to walk in appropriate directions for the location e.g. bi-directional flow in a corridor location
- Pedestrians tend to avoid other pedestrians and obstacles

We created stimuli of 3D virtual static pedestrian scenes, consisting of real formations extracted from photos, random formations, and synthetic formations based on our context rules described above. Our results showed that for an open, unconstrained location, participants judged scenes containing context formations to be as real as the real scenes, and more realistic than the random scenes. For constrained, corridor locations, people were able to distinguish between real and synthetic formations, but we found that scenes containing context rule based formations were judged to be almost as realistic as the real scenes. This implies that our context rule based pedestrian formations are a suitable alternative to manual placement.

3.2 Groups

Following on from this, we investigated the role of small groups in crowded pedestrian scenes [11]. In any urban pedestrian scene, it is assumed that a high percentage of the crowd will be in a group of two or more. Our first set of experiments on the role of groups in crowd plausibility investigate whether the addition of groups to a crowd scene makes it appear more realistic. We created a number of short video clips depicting scenes containing dynamic agents. These scenes contained either: no groups; plausible

Figure 2: Example of stimuli used in our dynamic group experiment [11], with no group scene on left, implausible group combination in centre and plausible group combination on the right.

group combinations; or implausible group formations (Figure 2). Plausible group formations contained higher numbers of individuals (singles) and groups of two (pairs) than groups of three (triples), while implausible group formations contained more triples than pairs or singles. We found that the addition of groups to dynamic crowd scenes improved the realism of such scenes, but only if the group sizes and numbers were plausible i.e., if there were more singles and pairs than triples.

Finally, we investigated the plausibility of conversational groups. These groups were created by motion capturing the movements, and recording the voices of three actors conducting a non-scripted conversation. While the conversations appear very realistic, the process is very time consuming. Therefore we conducted a series of experiments to determine methods to generate new conversations based on a finite data set, looking at body motion alone, and body motion and audio [14, 15]. We found that plausible and consistent conversational body motions are more important than matching audio and gestures. We also found that certain synthetic conversations can appear sufficiently realistic once talker and listener roles are correct, regardless of whether the body motions are temporally desynchronised or if the audio does not match the body motions.

4 System Implementation

In order to improve the realism of our crowd system, we added plausible conversational and dynamic groups. This work aims to build on the results found from our perceptual experiments, implementing them in our real-time crowd system to create a realistic crowd that behaves in a plausible manner (Figure 1). We use a tool that semi-automates the placement and orientation of characters that adhere to our context rules described in Section 3.1 [16]. This could be used, for example, to place characters for pre-visualisation and architectural visualisation purposes. Conversational agents have been implemented as a separate agent-type and can use either real conversation clips, or a selection of our synthetic conversations based on the results from our perceptual experiments.

Our dynamic behaviour system creates a population of individuals with entry and exit points and calculates each individual's path at runtime. In order to introduce groups to this system, we define three different types of agents: leaders, followers and neutrals. Neutrals behave as individuals and will compute and navigate individual paths around the environment. A leader is assigned a number of followers to select from their nearest neighbours, depending on the desired group size. As indicated from our experiments with dynamic groups, the majority of groups are of size two, with a small percentage containing more than two individuals. Once a leader has selected its followers, the follower's trajectory is altered to follow the leader and become part of the group.

5 Future Work

The next steps for the behaviour of our crowd system consists of two main stages. Firstly, we wish to add functionality to our agents to transition between static and dynamic agents. This is an important element, and will allow for realistic behaviours such as agents leaving and joining static groups and can reflect everyday behaviours such as people waiting, or meeting friends in a crowd. Further to this, we plan to add interactive functionality to our dynamic agents, to have groups appearing to be conversing as they navigate through the environment. Other group dynamics will have to be investigated also, such as group collision avoidance. For example, when is it appropriate for an entire group to change trajectory; and when is it more appropriate for a group to split up temporarily and merge back after an obstacle has been avoided?

Acknowledgements

We would like to thank Christopher Peters for his contribution to the work described in this paper, and Science Foundation Ireland for funding this work.

References

[1] J. Zacharias. Pedestrian behavior pedestrian behavior and perception in urban walking environments. *Journal of Planning Literature*, 16(1):3–18, 2001.

[2] W. Daamen and S. Hoogendoorn. Controlled experiments to derive walking behavior. *journal of Transport and Infrastructure Research*, 3(1):39 – 59, 2003.

[3] S.J. Older. Movement of pedestrians on footways in shopping streets. *Traffic Engineering and Control*, 10(4):160–163, 1968.

[4] E. Goffman. *Relations in Public: Microstudies of the Public Order*. Basic books, 1971.

[5] M. Wolff. *Notes on the behaviour of pedestrians*.

[6] R. Dunbar, N. Duncan, and D. Nettle. Size and structure of freely forming conversational groups. *Human nature*, 6(1):67–78, 1995.

[7] A. Lerner, Y. Chrysanthou, and L. Dani. Crowds by example. *Computer Graphics Forum*, 26(3):655–664, 2007.

[8] A. Lerner, E. Fitusi, Y. Chrysanthou, and D. Cohen-Or. Fitting behaviors to pedestrian simulations. In *Proceedings of ACM SIGGRAPH/Eurographics symposium on Computer animation*, pages 199–208. ACM, 2009.

[9] M. Neff, M. Kipp, I. Albrecht, and H.P. Seidel. Statistical reconstruction and animation of specific speakers gesturing styles. *ACM Transactions on Graphics*, 27(1):1–24, 2008.

[10] J. Cassell, H. Vihjalmsson, and T. Bickmore. Beat: the behavior expression animation toolkit. In *Proceedings of the 28th annual conference on Computer graphics and interactive techniques*, pages 477–486, 2001.

[11] C. Peters and C. Ennis. Modelling groups of plausible virtual pedestrians. *to appear IEEE Transactions on Computer Graphics and Applications*, 2009.

[12] C. Peters, C. Ennis, R. McDonnell, and C. O'Sullivan. Crowds in context: Evaluating the perceptual plausibility of pedestrian orientations. In *Proceedings of Eurographics Short Papers 2008*, pages 227–230, 2008.

[13] C. Ennis and C. Peters. Perceptual effects of scene context and viewpoint for virtual pedestrian crowds. In *To appear in Transactions on Applied Perception*, 2009.

[14] Rachel McDonnell, Cathy Ennis, Simon Dobbyn, and Carol O'Sullivan. Talking bodies: Sensitivity to desynchronization of conversations. *ACM Transactions on Applied Perception*, 6(4):22:1–22:8, 2009.

[15] Cathy Ennis, Rachel McDonnell, and Carol O'Sullivan. Seeing is believing: Body motion dominates in multisensory conversations. *To appear in ACM Transactions on Graphics*, 29(3), 2010.

[16] C. Peters and C. O'Sullivan. Metroped: A tool for supporting crowds of pedestrian ai's in urban environments. In *To appear in Proceedings of AISB Symposium on AI and Games Short Papers*, 2009.