Perceptual Evaluation of Position and Orientation Context Rules for Pedestrian Formations

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

Figure 1: *Scenes used in our experiment evaluating viewers' perceptions of Position & Orientation (left) and Position only (right)*

Abstract

In this paper, we evaluate the effects of position and orientation on the plausibility of pedestrian formations. In a perceptual study we investigated how humans perceive characteristics of virtual crowds in static scenes reconstructed from annotated still images where the orientations and positions of the individuals have been modified. We found that by applying rules based on the contextual information of the scene, such as the type of scene being portrayed, the presence of nearby individuals and objects and the constraints of the walking areas in the scene, we improved the perceived realism of the crowd formations. Results from this study can help in the creation of virtual crowds, such as computer graphics pedestrian models or architectural scenes.

CR Categories: I.2.10 [Artificial Intelligence]: Vision and Scene Understanding—Perceptual reasoning; I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism—Virtual reality

Keywords: perception, virtual crowd formation

1 Introduction

Humans are generally expert at recognizing and rating the behaviors related to other individuals. Given our high exposure to human behavior in our social environment, coupled with internal mechanisms, such as mirror neurons [Gallese et al. 1996] and those related to mentalizing and theory of mind [Premack and Woodruff 1978], we can discern the behavior of others at multiple levels of sophistication to varying degrees of detail and certainty. These capabilities are applicable too to robots and computer characters with a humanoid appearance [Schilbach et al. 2006], in which cases a viewer may rate the realism of an artificial motion or behavior favorably or unfavorably even if they are at a loss to identify precisely why it appears so. What is less clear however, and becoming evermore important as computer processing capabilities allow the display and simulation of increasingly large numbers of humanoid agents, is the question of the degree to which humans can discern and rate features of crowds of humans or humanoid figures.

In a previous pilot study focusing on pedestrian crowds in an urban context [Peters et al. 2008], we used static scenes consisting of crowds of humanoid characters to study the effects of orientation on users' perceptions of realism. The orientations of some characters were matched with those of their counterparts from real scenes, while others were artificially modified. We found that participants were consistently able to distinguish between the scenes containing the real orientations and those with artificial orientations. This initial indication that humans seem adept at judging such details raises intriguing questions as to what aspects are contributing to their impressions of realism and to what degree. In the current work, we take a significant step further by considering the effects of individuals' positions (in isolation and in combination with orientation modifications) on the perception of realism from a viewer (see Figure 1 for examples of images used for this evaluation). In addition, we introduce a proprietary tool to semi-automate the image generation process by incorporating contextual rules for the various formations, significant for eliminating manual errors and speeding up the process of creating the corpus to be used in the experiment.

We hypothesize that the context in which each individual is perceived is one, and perhaps the most important, of these aspects. By *context*, we refer to each individual's relationship with respect to both its environment and neighbors. These relationships should be appropriate and consistent with common experience: pedestrians tend to be seen walking on paths, are often directed towards exit or goal positions when mobile, do not walk into obstacles or other individuals, and members of groups are usually in proximity to each other and may have other properties in common. In order

*e-mail:[ennisca, Christopher.Peters, Carol.OSullivan]@cs.tcd.ie

APGV 2008, Los Angeles, California, August 9–10, 2008.
© 2008 ACM 978-1-59593-981-4/08/0008 $5.00

to test this, we compare scenes where individuals' positions and orientations have been generated according to contextual rules with those directly derived from real scenes, and those that have been randomly generated. It is important to remember that the objective of this study was not to physically replicate real scenes, but to generate formations of pedestrians in a way that is *perceptually* real to the viewer by considering the context of the scene. This work is particularly important for evaluating and modeling pedestrian and crowd behavior (see [Shao and Terzopoulos 2005] and [Pelechano et al. 2007] for recent examples).

After reviewing related work in Section 2, Section 3 outlines the methodology for constructing the scenes used in our experiment. We detail the perceptual experiment in Section 4 and describe their results and possible applications in Section 5.

2 Related Work

While there has not been much research focusing on the perception of crowds, it is still an important factor and there are many methods of simulating crowds and crowd behavior. Many different approaches are taken, such as using social force models [Helbing and Molnár 1995], path planning [Lamarche and Donikian 2004] or behavioral models incorporating perception and learning [Shao and Terzopoulos 2005] or sociological effects [Musse and Thalmann 1997].

In recent years, human perception of virtual worlds, their characters and associated animations has been increasingly recognized as an important factor in achieving more realistic scenes. While much research has been conducted regarding perception of animation and motion of individuals (for example [Reitsma and Pollard 2003], [McDonnell et al. 2007]), or spatial awareness [Henry and Furness 1993] very little has been studied in relation to crowds, in particular crowd formation.

There are two previous approaches that are similar to our work, particularly in terms of data collection and annotation. Lerner et al. [2007] set out to generate crowds that displayed varied behaviors individually without defining an explicit behavior model. They used a data-driven example based approach to achieve this, allowing the agents to learn from real-world examples. A database is constructed from input video of real world pedestrian behaviors by manually tracking pedestrians in the video to generate a set of trajectories, which are stored as examples in the database. At runtime, the database is queried for similar examples that match those of the simulated pedestrian and the closest matching example is selected as the resulting trajectory.

Lee et al. [2007] use a data-driven approach to simulate virtual human crowds imitating real crowd behavior. They recorded crowd videos in a controlled environment from an aerial view. Users must manually annotate video frames with static environment features and can semi-automatically track multiple individuals in order to provide their trajectories. This data informs an agent movement model to provide a crowd that behaves similarly to those observed in the video. Although both cases involve manual annotation of crowd behavior, both use the data as exemplars for generating behaviors rather than as a basis for conducting perception studies.

In our previous study [Peters et al. 2008], we examined the perceptual plausibility of pedestrian orientations in isolation. It was found that participants were able to consistently distinguish between those virtual scenes where the character orientations matched the orientations of the humans in the corresponding real scenes and those where the character orientations were artificially generated, according to a number of different rule types. In addition, orientation rule types that accounted for the context of the scene were judged overall to be more realistic than those utilizing naive rule-sets.

Figure 2: *Two reconstructed scenes from real scene still images. (a) is the corridor zone with 12 pedestrians. (b) the open zone with 30 pedestrians*

3 Methodology

Our methodology consists of four phases. The first three phases refer to the collection (Section 3.1), annotation (Section 3.2) and reconstruction (Section 3.3) of virtual scenes closely approximating original still images in terms of pedestrian positions and orientations. The final phase is the modification (Section 3.4) of aspects of these scenes according to rules, in order to produce artificial formations to be compared with each other and the real reconstructions.

3.1 Data Collection Phase

A number of videos were taken of two different locations, each representing an archetypical pedestrian movement zone. We refer to these as *constrained* or *corridor* zones and *unconstrained* or *open* zones (see Figure 2). An open zone represents a relatively large space where pedestrians tend to be seen crossing in many varying directions due to the presence of many possible exits and entrances. In contrast, a corridor zone is more constrained, usually with a single entrance/exit at either end and therefore tends to enforce bidirectional movement.

A number of still images were extracted from each video, to be used as a basis from which to create reconstructions of the scenes depicting the real positions and orientations of individuals. These will be referred to here as the *real* category of scenes. To minimize the variation in responses from participants, the density of pedestrians in the two zones (corridor and open) was kept as equal as possible. Based on the area of the zone visible to the viewer, it was calculated that 30 pedestrians in the open zone would roughly correspond in crowd density terms to 12 in the corridor zone, and therefore still images with these numbers were selected from the extracted video stills for the respective zone types.

3.2 Annotation Phase

Each still image was annotated manually to highlight individuals' positions and orientations and their groupings, if any (see Figure 3(b)). Groups were deemed to consist of one or more individuals, according to their localization in space and aided by a visual inspection of the video clip surrounding the still image being annotated. Each group was designated by an ellipse, which covered all members of the group and was color-coded according to whether the corresponding group was static (*black*) or mobile (*yellow*). The orientations of individuals were classified as belonging to one of the following 8 rotations specifying cardinal directions: 0, 45, 90, 135, 180, 225, 270 and 315. Each direction was associated with a

Figure 3: *Reconstructing the scene. An (a) initial still image is (b) annotated with groups and their orientations. Camera parameters are matched up with those of the original camera so (c) characters can be placed corresponding to the transformations of the real people resulting in (d) a virtual scene with a similar composition to the real one.*

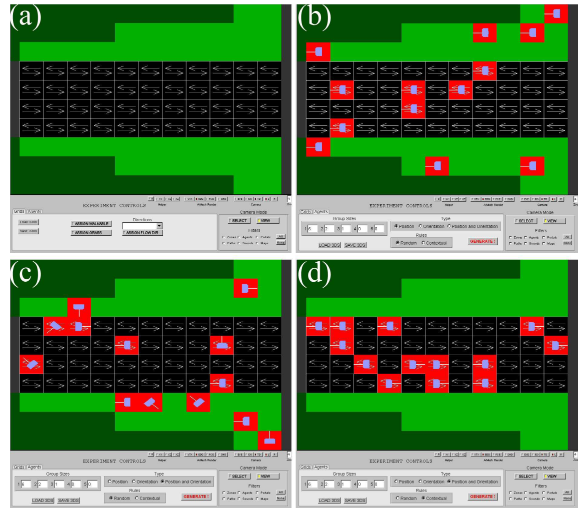

Figure 4: *View of our proprietary tool for generating pedestrian transformations based on contextual and random rules, in this case for the corridor scene: (a) design of the grid with out-of-bounds and flow directions; (b) examples of output for random position; (c) random position and orientation and (d) context modified position and orientation rules.*

unique color code, to aid visual recognition of the general characteristics of the scene, such as the number of groups containing one, two or three individuals.

3.3 Reconstruction Phase

The reconstruction phase consists of recreating virtual replicas of the real images captured and annotated in the previous phases. Once a still image has been annotated, we reconstruct the scene by using it as a viewport background in *3D Studio Max* and fitting our 3D model by manually tweaking the virtual camera parameters to ensure an acceptable fit between the still image and the model. Next, the positions of virtual characters are manually matched up with their real-life counterparts from the still image, providing a good approximation to the composition of the original scene.

These reconstructions were replicas in the sense that they copied certain aspects from the real scenes, such as individuals' positions, orientations and groupings, whereas we did not attempt to replicate individuals' appearance, clothes and gender.

3.4 Modification Phase

We investigated pedestrian formations according to which of the following three aspects would be studied: only individuals' positions, only their orientations and both their positions and orientations. For each, we generated the required positions and/or orientations semi-automatically, using a combination of our own standalone proprietary tool (see Figures 4 and 5) and 3D Studio Max. The proprietary tool was important not only for reducing the workload in creating the scenes to be used in the experiment, but also for reducing placement errors and aiding in the replicability of the experiment in general.

Since an important focus for us is to consider the *context* of individuals the scene, rather than looking at the pedestrian characteristics in isolation, a number of context rules were defined for positioning and orientating characters. There can be many different aspects relating to context, which belong to three general types: firstly, nearby pedestrians, objects and obstacles that may affect an individual; secondly, the type of walking area that an individual inhabits, e.g., in order to specify the general direction of *flow* in that area [Chenney 2004]. Finally, groups may play a large role in people's perceptions of crowds and pedestrians e.g., group size and the number of groups in a scene.

For those rules modifying the original positions or orientations, a number of steps must be taken in our proprietary tool to allow for the automatic generation of data. A grid is created in order to fit the area that will appear in the final scene. Each cell in the grid is then manually assigned with attributes, such as walkability and flow direction(s), if any. This process only needs to be conducted once for each zone from which renderings must be conducted: in our study, this process needed to be conducted only twice, once for the open scene and once for the corridor scene. One can then select a position type and a rule to apply - clicking on a button will then generate the resulting transformations according to the rules selected.

After a scene has been generated with the tool, it is exported to 3D Studio Max as a set of dummy nodes, each of which contains the transformation for a particular pedestrian. Each node is manually associated with a mesh, either a posed human figure or else a direction-less pawn figure, the latter of which is used in the position studies.

The rules and steps involved in the modification of the scenes are described in Section 3.4.1 for position, 3.4.2 for orientation, and Section 3.4.3 for orientation and position.

3.4.1 Position Rules

For the images modified for the position formations, the characters are displayed as *pawn* characters with no discernible orientation. The images used for the position block of the experiment are shown in Figure 6. Here, the only information available is the position of the characters, and this is modified in the following ways:

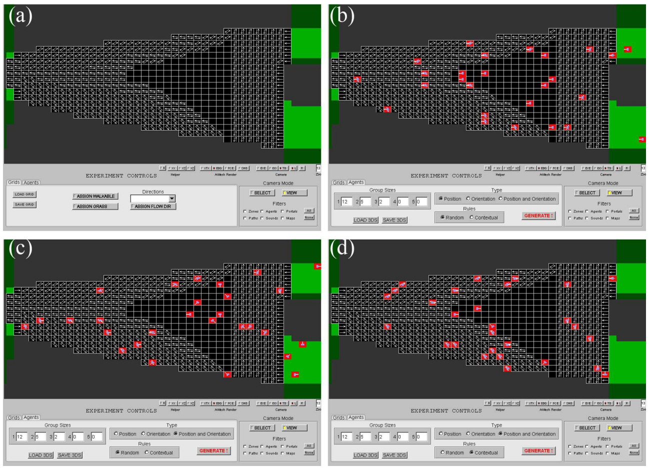

Figure 5: *Generation of pedestrian transformations for the open scene: (a) design of the grid with out-of-bounds and flow directions; (b) examples of output for random position (c) random position and orientation and (d) context modified position and orientation rules.*

1. Real - The position of each individual in the scene is the same as the positions of the pedestrians in the still image.
2. Random - Each individual is assigned a random position on the grid.
3. Context based - Each individual is assigned a position according to our context rules, which are listed below.

Position Context Rules

1. Bounds Sensitive - An individual can only be assigned a position that is part of a designated walkable area. In these experiments, grass was regarded as being *out-of-bounds* during the application of contextual rules.
2. Group Sensitive - Individuals will be assigned a position to maintain the number and size of groups in the original still image. In other words, in the case of a group of pedestrians, the position of the group only will change, not the position of the individuals making up the group.

3.4.2 Orientation Rules

The modified images used for the orientation block contained standard human character models with discernible orientations (see Figure 7). The character positions remained the same as those from the original still image; only their orientations were changed.:

1. Real - The orientation of each individual in the scene is the same as the orientations of the pedestrians in the still image.
2. Random - Each individual is assigned one of the 8 cardinal orientations on a random basis.
3. Context - Each individual is first assigned a random orientation, which then is altered to obey our context rules (listed below).

The orientation context rule can be further specified as consisting of the application of the following three rules when determining the orientation of each individual:

1. Flow Sensitive - the orientation of each individual is chosen randomly from a subset of the 8 cardinal directions. This subset is created from the allowable flow directions for the position of the character, based on a ground flow-tile representation.
2. Adjacency Sensitive - the orientation of each individual is chosen randomly from the 8 cardinal directions, but any direction leading to inappropriate facings is disqualified e.g. a character walking into a lamp post would be considered an inappropriate facing.
3. Group Sensitive - Each individual is assigned orientations according to their group membership, rather than on an individual basis, if they were part of a group of 2 or more.

Figure 6: *Position images for the open scene: (a) the original still image; (b) the virtual representation of the real still image; (c) the open scene with random positions and (d) the open scene with position context rules*

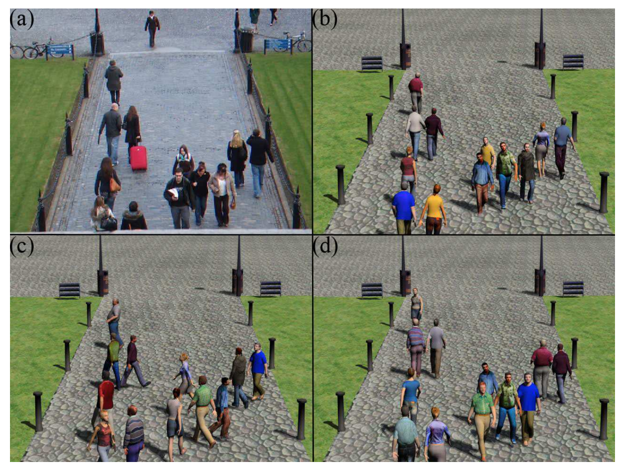

Figure 7: *Orientation images for the corridor scene: (a) the original still image; (b) the virtual representation of the real still image; (c) the corridor scene with random orientations and (d) the corridor scene with orientation context rules*

In order to resolve conflicts between the rules, the rules were applied in the order provided above: flow sensitivity, then adjacency sensitivity and, if applicable, group sensitivity. Members of a group were all orientated in the same direction, based on the direction in which most of the group members were facing in accordance with the previously applied rules. The scenes containing orientation were constructed manually.

Figure 8: *Position and orientation images for the open scene: (a) is the original position and orientation; (b) position context rules with random orientation; (c) random position and orientation context rules and (d) both position and context rules*

3.4.3 Orientation and Position Rules

For the images modified for the orientation and position formations, the scene is again divided into a grid. As in Section 3.4.2, this block of the experiment used normal human characters with discernible orientations. The context rule images used for this block are shown in Figure 8. Both the orientation and position will be modified in the following ways:

1. Real - The position and orientation of each individual in the scene is the same as the positions and orientations of the pedestrians in the still image.
2. Random - Each individual is assigned a random position on the grid and orientation from the 8 cardinal orientations.
3. Position Context (Context Pos) - Each individual is assigned a random orientation with a position according to the position context rules as explained in Section 3.4.1.
4. Orientation Context (Context Ori) - Each individual is assigned a random position on the grid and an orientation according to the orientation context rules as explained in Section 3.4.2.
5. Orientation and Position Context (Context Both) - Each individual is assigned an orientation and position that obey both our orientation and position context rules.

4 Experiment

Thirty two participants (12F, 20M) age 18 to 30, were seated in front of a computer screen. They were told that the experiment consists of three blocks and were given an instruction sheet: two photographs of the corridor and open zone were shown and they were told that the images they were about to see were derived from real photographs, but in some the character formations were real, while in others they were synthetically generated. For the first block of the experiment the participants were told to focus only on the positions of the characters. For each image displayed, participants were asked if they thought the positions of the *pawn* figure characters were real or synthetically generated. For the second block, participants were asked to look at the orientations of the characters only and judge if they were real or synthetically generated. For the final block of the experiment, participants were asked to take both position and orientation of the characters into account and judge whether the scenes were real or synthetically generated. The reason that we presented the experiment in this order was to avoid biasing participants. If the pawn figures were viewed after the humanoid characters, this could have caused them to perceive the scenes as less realistic due to the reduced realism of the characters, which was not the effect being tested. Furthermore, the scenes with position and orientation combined were presented during the final block, to prevent participants from taking position into consideration when conducting the orientation only trial. Between each trial, a blank-screen was displayed for 5 seconds, after which the number of the next trial was displayed alerting participants.

For the first block of the experiment, a total of 24 images were displayed for 4 seconds each, 12 of which were master scenes where the positions matched an original still image, and 12 of which were positions modified by the rules in Section 3.4.1. An image could thus be categorized as belonging to one of the following three different types: Original, Random and Context.

For the orientation block of the experiment, again 24 images were shown for 4 seconds each, 12 containing scenes with real orientations and 12 containing modified orientations modified as described in Section 3.4.2 with Original, Random and Context categories.

For the final experiment block, looking at position and orientation, a total of 48 images were displayed for 4 seconds each. Of these, 24 contained positions and orientations matching still images of real scenes, and 24 images were modified according to the rules mentioned in Section 3.4.3. For this block of the experiment, an image could be categorized as belonging to one of the following five different types: Original, Random, Position Context, Orientation Context and Both Context.

5 Discussion

We found no significant differences based on the ordering of the images within the experiment blocks in the responses of the participants, implying that there were no ordering effects within the individual experiments.

5.1 Position Results

We averaged responses over each of the three repetitions for each position type. A two factor ANOVA with repeated measures showed a main effect of Scene ($F(1, 31) = 17.895, p < 0002$), in that the open scene was perceived to be more realistic than the corridor scene, and Position type ($F(2, 62) = 70.077, p < 0.00001$) where the real positions were perceived to be more real than virtual positions. There was also interaction between the two ($F(2, 62) = 23.476, p < 0.00001$), where the real positions were perceived as real more often for the corridor location than the open location but the virtual positions were both perceived as real more often for the open location. Post-hoc analysis was then performed using a standard Newman-Keuls test for pairwise comparisons among means.

We found that the original corridor scenes with real positions were judged as *real* significantly more times than either the context or random virtual corridor scenes ($p < 0.0002$ in all cases), implying that participants are able to distinguish the real cases from the synthetic ones based primarily on differences in position for constrained zones. However, participants perceived the scenes with context ruled positions to be more real than those with random positions ($p < 0.05$).

Looking at the open scene, while participants perceived the real positions more *real* than the random positions ($p < 0.0002$), they

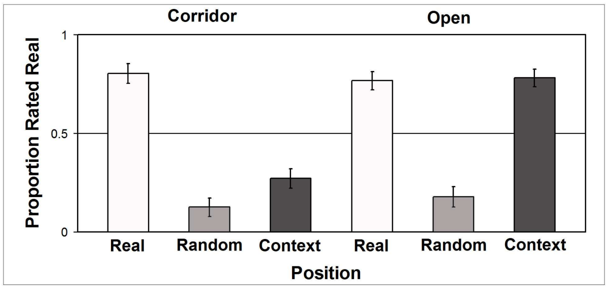

Figure 9: *Graphs showing position results for Corridor and Open Scenes*

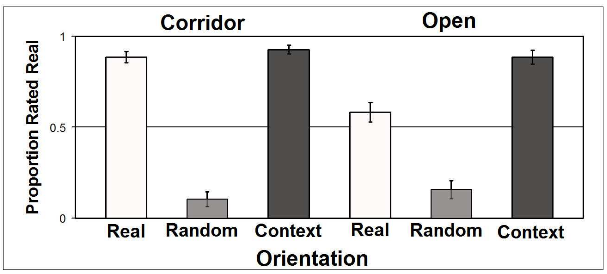

Figure 10: *Graphs showing orientation results for Corridor and Open Scenes*

judged the scenes with context-based positions to be as realistic as those with the real positions. Figure 9 shows these findings on a scale of 0 to 1, where 0 means they were perceived as synthetically generated and 1 means they were perceived as real.

As can be seen from Figure 9, while the context rules applied did improve the perceived realism of the scene, the participants could still distinguish the real positions from the synthetically generated ones. For for the open scene, the context rules applied had a greater effect on participants perceptions of the scene. The fact that participants judged the scenes with our context rules to be as real as those scenes with real positions indicates that these rules would be an adequate way to populate scenes when positioning characters in an open or unconstrained zone. While the rules do not have such a strong effect on the realism of corridor scenes, it has been shown here that they would be a suitable alternative to random positioning for characters, while not necessitating manual placement of individuals. While this is interesting in itself, the effects of the positioning of characters are not very useful for practical applications without taking into account the orientations of characters. This will be discussed further later in this section.

5.2 Orientation Results

We averaged responses over each of the three repetitions for each orientation type. A two factor ANOVA with repeated measures showed a main effect of Scene ($F(1, 31) = 11.508, p < 0.002$), where the Corridor scene was perceived to be more real for this experiment, and Orientation type ($F(2, 62) = 162.04, p < 0.0001$) where context orientations were perceived to be the most real, followed by the real orientations, with the random orientations being judged the least real. There was also interaction between Scene and Orientation ($F(2, 62) = 12.040, p < 0.00005$), where participants judged the real and context scenes as real more often for the corridor location but judged the random scene as real more often for the open location. Post-hoc analysis was then performed using a standard Newman-Keuls test for pairwise comparisons among means.

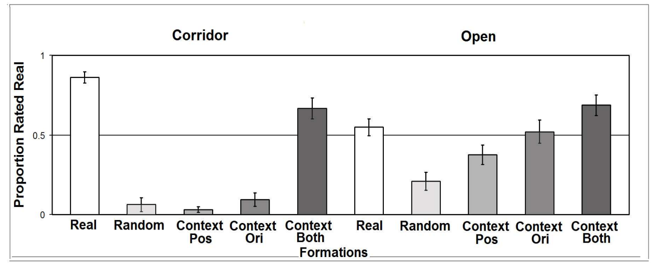

Figure 11: *Graphs showing position and orientation results for Corridor and Open Scenes*

We found that the original corridor scenes with real orientations were judged as *real* significantly more times than the random virtual corridor scenes ($p < 0.0002$), but the scenes with context rules were judged to be as real as the original scenes, implying that participants are no longer able to distinguish the real cases from orientations generated using our context rules for constrained zones.

Looking at the open scene (see Figure 10), while participants perceived both the real and context-based orientations to be more *real* than the random orientations ($p < 0.0002$), they judged the scenes with context-based orientations to be *more* realistic than those with the real orientations. One explanation for this interesting result is that the unconstrained nature of the open scene provided less contextual cues with which to judge the validity of the characters' orientations. This would suggest that there may be more perceptual tolerance when viewing character formations in these scene types than in more constrained types, where viewers are more adept at spotting peculiarities.

As illustrated in Figure 10, the results show that the context rules greatly affect how real the orientations are perceived to be. In our previous study, our results showed similar trends. However, the addition of the adjacency sensitive rules to the context scenes resulted in the participants perceiving the orientations as significantly more realistic than in our previous study, in particular for the open scene. This is expected due to the area of the open scene, particularly in the middle of the scene, where the adjacency rules are applied. It's likely that the participants focus more on the center of the image for the duration of exposure to the image, explaining the positive effect on perceived realism caused by the addition of the adjacency rules. As with the results for the position rules, the combination of position and orientation rules are where the most useful information is in terms of practical applications for placing human characters in virtual scenes.

5.3 Position & Orientation Results

We averaged responses over each of the three repetitions for each type of formation. A two factor ANOVA with repeated measures showed a main effect of Scene ($F(1, 31) = 15.754, p < 0.0005$), where the open scene was perceived to be more real than the corridor scene, and formation type ($F(4, 124) = 54.093, p < 0.0001$), where scenes with both position and orientation context rules were judged to be almost as *real* as the real scenes. There was also interaction between the two ($F(4, 124) = 16.615, p < .00001$) where participants perceived the real formations to be more real for the corridor location, but judged each virtual formation to be more real for the open location. Post-hoc analysis was then performed using a standard Newman-Keuls test for pairwise comparisons among means.

We found that for the corridor scenes, the real formations were

judged as *real* significantly more times than any of the four virtual formations ($p < 0.02$ in all cases). This implies that, given position and orientation information for a constrained scene, participants are able to tell real scenes from virtual scenes. Having said that, the scenes where both position and orientation context rules had been applied were perceived as real significantly more often than either the random, position context or orientation context scenes ($p < 0.0004$ in all cases).

In the open scene, the results were again more complicated, with the real scenes being judged as real more often than the random formations and scenes with position context information only ($p < 0.004$ in all cases). For the scenes with orientation context and both position and orientation context, the participants judged them to be as real as the real scenes. This reiterates the possibility that participants find it more difficult to distinguish between real and synthetic formations for less constrained scenes.

From Figure 11, it can be seen that, when context is considered when placing characters, participants perception of the realism of the characters is greatly improved compared to random placement. While participants can still differentiate between real and synthetic corridor scenes, they cannot differentiate between these for open scenes. Interestingly, for the corridor scene, the use of either position or orientation context on its own was not effective, as participants perceived these to be as synthetically generated as the random scenes. However, for the open scene, participants perceived the scenes with random positioning and orientation context rules to be as real as both the real scenes and the scenes with both position and orientation context rules applied. This would imply that, in an unconstrained scene with a large number of people, orientation seems to be of greater importance when it comes to plausibility.

6 Conclusions & Future Work

Overall, for the open location, participants judged those scenes utilizing our contextual rules to be as real, or more real, than the real scenes. This can be explained by the fact that due to the large number of entrances and exits and the space available, a wider variety of formations were acceptable as being realistic since less constraints could be applied by the viewer.

In general, for the corridor location, participants were better able to discern between the real and artificial, due to the constrained nature of the zone. Participants were particularly sensitive to the positioning of figures in this zone, something that needs to be considered in more detail in future experiments. Despite this, overall, our combined contextual rules for this zone were judged to be nearly as realistic as the real scenes.

The fact that most of the participants were familiar with the area used for both scenes, in particular the open scene, means that they were aware of the general directions of flows of pedestrians and this is reflected in their judgements of realism throughout the experiment. For the orientation and positions and orientation combined experiments, participants judged the real images real 58% and 55% of the time respectively. This could be explained by the fact that in a real setting, there will be a certain amount of randomness to the scene e.g. a character moving in a direction not usually taken in a flow lane. For both of these settings, participants rated the images with context-rules as real more often (89% and 69% respectively), which implies that when people are familiar with an area, they tend to look for these context elements in a static scene when making judgements on realism.

To summarize, our results show that a viewers ability to distinguish between real and artificial scenes depends on the context of the scene and that contextual factors are vital when considering the perceived realism of pedestrian formations. From a modeling standpoint, the results from these experiments, most significantly those taking both position and orientation into account, imply that the contextual rules presented here form an effective general starting point from which to populate urban environments. Applications of this work include creating initial formations of characters for populating real-time virtual environments and the placement of pedestrians in static urban architectural displays. There are also longer-term possibilities for application to level-of-detail metrics for pedestrian and crowd behaviour models (c.f. [Shao and Terzopoulos 2005] and [Pelechano et al. 2007]), where accuracy of simulation may be traded for performance in order to display larger crowds than would otherwise be possible while minimizing perceived errors in the behaviour.

Nevertheless, many improvements and perceptual investigations remain to be made. Of particular importance is a more detailed investigation of factors at the group level e.g. group sizes, number and distribution of groups and group formations under both static and dynamic conditions. It seems likely that groups play an important role in the perception of pedestrian scenes. No doubt there are a vast array of factors for possible consideration when creating context rules for groups: it will be challenging to identify, isolate and choose a subset of these in order to be able to conduct tractable experiments.

We also plan to consider dynamic pedestrian scenarios and study the formulation of dynamic context rules derived from those described here for static situations. Perceptual evaluation of such rules and comparisons with existing models e.g. those based on steering behaviours [Reynolds 1987], could be of great utility in the evaluation and construction of crowd simulations for interactive applications.

7 Acknowledgements

The authors wish to thank the reviewers for their constructive and helpful feedback.

References

CHENNEY, S. 2004. Flow tiles. In *SCA '04: Proceedings of the 2004 ACM SIGGRAPH/Eurographics symposium on Computer animation*, Eurographics Association, Aire-la-Ville, Switzerland, Switzerland, 233–242.

GALLESE, V., FADIGA, L., FOGASSI, L., AND RIZZOLATTI, G. 1996. Action recognition in the premotor cortex. *Brain 119*, 593 – 609.

HELBING, D., AND MOLNÁR, P. 1995. Social force model for pedestrian dynamics. *Phys. Rev. E 51*, 5 (May), 4282–4286.

HENRY, D., AND FURNESS, T. 1993. Spatial perception in virtual environments: Evaluating an architectural application. *Virtual Reality Annual International Symposium, 1993., 1993 IEEE* (Sep), 33–40.

LAMARCHE, F., AND DONIKIAN, S. 2004. Crowd of virtual humans: a new approach for real time navigation in complex and structured environments. *Computer Graphics Forum 51*, 3 (Sept), 509–518.

LEE, K. H., CHOI, M. G., HONG, Q., AND LEE, J. 2007. Group behavior from video: a data-driven approach to crowd simulation. In *SCA '07: ACM SIGGRAPH/Eurographics symposium on Computer animation*, 109–118.

LERNER, A., CHRYSANTHOU, Y., AND DANI, L. 2007. Crowds by example. *Computer Graphics Forum 26*, 3, 655–664.

MCDONNELL, R., NEWELL, F., AND O'SULLIVAN, C. 2007. Smooth movers: perceptually guided human motion simulation. In *SCA '07: Proceedings of the 2007 ACM SIGGRAPH/Eurographics symposium on Computer animation*, Eurographics Association, Aire-la-Ville, Switzerland, Switzerland, 259–269.

MUSSE, S. R., AND THALMANN, D. 1997. A model of human crowd behavior: Group inter-relationship and collision detection analysis. In *Proceedings of the 1997 Eurographics Workshop on Animation and Simulation*, 39–52.

PELECHANO, N., ALLBECK, J. M., AND BADLER, N. I. 2007. Controlling individual agents in high-density crowd simulation. In *SCA '07: Proceedings of the 2007 ACM SIGGRAPH/Eurographics symposium on Computer animation*, Eurographics Association, Aire-la-Ville, Switzerland, Switzerland, 99–108.

PETERS, C., ENNIS, C., MCDONNELL, R., AND O'SULLIVAN, C. 2008. Crowds in context: Evaluating the perceptual plausibility of pedestrian orientations. In *Short Papers Proceedings of Eurographics 2008*.

PREMACK, D., AND WOODRUFF, G. 1978. Does the chimpanzee have a theory of mind? *Behavioural and Brain Sciences 1*, 515 – 526.

REITSMA, P. S. A., AND POLLARD, N. S. 2003. Perceptual metrics for character animation: sensitivity to errors in ballistic motion. *ACM Trans. Graph. 22*, 3, 537–542.

REYNOLDS, C. W. 1987. Flocks, herds, and schools: A distributed behavioral model. *Computer Graphics 21*, 4, 25–34.

SCHILBACH, L., WOHLSCHLAEGER, A., KRAEMER, N., NEWEN, A., SHAH, N., FINK, G., AND VOGELEY, K. 2006. Being with virtual others: Neural correlates of social interaction. *Neuropsychologia 44*, 5, 718 – 730.

SHAO, W., AND TERZOPOULOS, D. 2005. Autonomous pedestrians. In *SCA '05: Proceedings of the 2005 ACM SIGGRAPH/Eurographics symposium on Computer animation*, ACM, New York, NY, USA, 19–28.