Perceiving Motion Transitions in Pedestrian Crowds

Qin Gu Chang Yun Zhigang Deng *
Department of Computer Science, University of Houston

Abstract

Creating natural motion transitions between different motion clips is crucial for reusing and editing character animations. Perception of motion transitions in a pedestrian crowd is affected by many collective features such as crowd density, appearance variations, motion variations, and sub-group interaction patterns. In this paper, we conducted a series of psychophysical experiments to investigate how these crowd features can influence human perception on walking motion transitions in a crowd when inexpensive motion blending algorithms are used. Our results provide useful implications and practical guidelines for performance-oriented crowd applications such as real-time games to improve the perceptual realism by effectively disguising motion transitions.

CR Categories: H.1.2 [Information Systems]: Models and Principles—User/Machine Systems - Human factors; I.3.7 [Computing Methodologies]: Computer Graphics—Three-Dimensional Graphics and Realism - Animation

Keywords: Crowd Simulation, Motion Transition, Motion Blending, Psychophysical Study, Perceptual Realism

1 Introduction

Crowd animations have been increasingly used in many entertainment and simulation applications in recent years. Besides the detailed animation of each crowd agent, the perceptual realism of a crowd is influenced by many "macro" features such as its density and variety. For instance, McDonnell and her colleagues thoroughly evaluated the perceptual influences of agent/character appearance diversity and motion diversity in a crowd [McDonnell et al. 2008; McDonnell et al. 2009]. Such work served as a useful guidance for fast real-time crowd simulations.

As one of the most common motions in crowd simulations, walking motion is especially important for simulating both casual scenarios such as metropolitan streets and collective behaviors such as military march. In a real-world crowd, the walking motion of each pedestrian is distinctive, with different gait styles [Cutting and Kozlowski 1977]. Even the same pedestrian may changes his/her walking motion patterns over time, such as looking in different directions or increasing/decreasing the pace step. In other words, the walking motions of a crowd typically vary in both spatial and temporal domains. However, it is time and memory consuming to create complete motion sequences for every pedestrian. A widely used approach for real-time crowd simulations is to pre-define a pool of different short walking clips by using 3D animation tools or motion capture, and then plausibly concatenate them together to mimic the subtlety of crowd motion variations at runtime.

While the above approaches help to save manual efforts in creating walking sequences for crowds, they also introduce a large number of motion transitions automatically blended among different keyposes or animation segments. Typically, these motion transitions are generated by simple linear or spherical interpolations which does not consider the subtle motion intermittence of real human motion frame by frame so that they look less realistic than motion capture data. In addition, audiences tend to form perceptual dissatisfaction on the transitions when the transition target motion appears quite different from the source one. Numerous efforts have been made to improve the visual quality of these motion transitions in the case of single character animations. For example, Kovar et al. [2002] proposed a novel graph structure to extract and connect the optimal motion transitions among motion segments. Spatial, temporal and inverse kinematic constraints were also used in synthesizing motion transitions to achieve physically plausible results [Rose et al. 1996]. To evaluate those synthesized motions for a single character, several studies have also been conducted to investigate human perception on unnatural motion transitions [Reitsma and Pollard 2003; van Basten and Egges 2009]. In theory, any optimization approach derived from single character animations can also be applied to individual agents in a crowd. However, the performance bottleneck of crowd simulations often requires animation practitioners to take advantage of some unique features of a crowd such as density and variety, to make an illusion that the animation quality of each character in the crowd is close to its single character counterpart without utilizing sophisticated optimization techniques.

The primary objective of our study is to investigate how "macro" crowd features such as crowd density can influence human perception on walking motion transitions in a crowd when inexpensive motion blending algorithms are used. Specifically, we perform a series of psychophysical experiments, described in Section 2, to evaluate the influences of *viewpoints, agent densities, appearance variations, motion variations* and *sub-group interactions* in pedestrian crowds. As a result (detailed in Section 3), we found most of them have notable effects on viewer perceptions, and deserve considerations when simulating a crowd. Finally, we discuss the assumption in the current study and future work (Section 4).

2 Experiments

Prior to the experiments, we generated different walking motion clips based on motion capture data. We selected eight mixed motion sequences with different walking styles (total 4334 frames) from the publicly available, CMU motion capture database (http://mocap.cs.cmu.edu) and trimmed them into short walking segments so that every segment only includes one full walking cycle. This provided a total of 16 units of walking cycles. Then, we performed a K-mean clustering on character joint angle space(Totally 42 DOF) to remove the motions too distant from the cluster centers so that no transition between any motion pair is dominantly observable. The final motion pool contains 10 walking motions as the runtime transition candidates. All the animation clips used in this work can be accessed at `http://graphics.cs.uh.edu/projects/CrowdTransitionPerception/`.

2.1 Specifications

The high-level crowd simulation model used in our experiments is the HiDAC model [Pelechano et al. 2007]. Five common "macro" features of simulated crowds are investigated in our psychophysical experiments, including viewing angle, crowd density, agent appearance variations, agent motion variations, and group interactions. A total of 36 students whose ages range between 18 and 33 with little crowd animation background participated in our experiments.

*e-mails:{qgu4|cyun|zdeng4}@uh.edu

First, we showed a brief document to participants to introduce the purpose of our experiments and requested them to focus on detecting changes in walking motions rather than modeling or rendering realism. They were asked to rate two randomly selected test clips to get familiarized with the experiments. This was done to reduce the rating inconsistency during the follow-up formal experiments. The generation rules of the motion transitions in all trials were kept the same, that is, each pedestrian randomly selected one of the 10 walking motion candidates as interpolation target after finishing one walking cycle. We applied a spherical linear interpolation on the agent rotation and a linear interpolation on the agent translation using 7 frames window to generate those transitions. In order to evaluate how well the participants can detect motion transitions, we informed them that different transition frequencies in different scenarios may exist, and asked them to rate their general perception of average transition numbers in the scale from 1 (no transition) to 10 (transition at every walking cycle).

Figure 1: *Left: Crowd density test scene(the strategy view, agent density: 64), Middle: Crowd appearance test scene (the FPS (First Person Shooter) view, texture styles: 16), Right: Crowd motion test scene (the FPS view, walking motion candidates: 10)*

Viewpoint is a parameter consisting of a viewing angle and a viewing distance. Based on the assumption that different viewpoints potentially affect the perception of crowd motions, we choose two common viewpoints in our experiments, which are usually observed in Real-time Strategy games (called "strategy view" in this work, left of Figure 1) and First Person Shooter games (called "FPS view" in this work, middle and right of Figure 1). For each viewpoint, we conduct 5 different agent densities + 5 different appearance numbers + 5 different motion numbers + 4 group interactions = 19 trials per each viewpoint. Therefore, the total number of trials for each participant is 38. The duration of each trial is experimentally set to 20 seconds. When testing each crowd feature, the other features were fixed to their default values (Table 1) which are the medians of the value ranges. The trial order was randomly generated for each participant to counter balance the experiments.

Crowd Features	Density	Appearance	Motion	Interaction
FPS/Strategy view	16	4	6	random

Table 1: *Default values of the studied crowd features*

2.2 Crowd Density Effect

The number of crowd agents is one of the most important features of crowd animations. It directly affects the visual realism and overall performance of the animations. Intuitively, the first parameter we controlled to test is the variation of agent numbers.

2.2.1 Experiment Description

When computing crowd density, the total number of agents in the crowd becomes inaccurate if the scale of environments or agent distributions vary significantly. Thus, in this study, we counted the average number of agents in the viewport as the density of a crowd. To verify whether the precise motion transitions in high density pedestrians are as observable as those in low density pedestrians, we evaluated the participants' perception data to the average motion transition numbers from a low average density of 4 to a high average density of 64 (the first row of Table 2) for both viewpoints (left of Figure 1).

Features	trial 1	trial 2	trial 3	trial 4	trial 5
Density	4	8	16	32	64
Appearance	1	2	4	8	16
Motions	2	4	6	8	10
Interactions	random	chase	flocking	advection	no interaction

Table 2: *Experiment Specifications*

2.2.2 Results

A two-way analysis of variance was used to evaluate the average transition frequencies rated by the participants. We found both the density of the crowd ($F = 12.89, p < 0.017$) and the viewpoint ($F = 32.91, p < 0.001$) were the main effects. Based on the average rating score, the first part of the result indicates that the number of pedestrians in the simulated scene affected the participants' perceptive ability to detect the motion transitions. Increasing the number of crowd agents, to a certain extent, disguises unsmooth motion changes from viewers. The second main effect of the result shows that the strategy viewpoint was more successful in hiding the motion transitions than the FPS viewpoint. One possible reason is that the viewing distance in the strategy view is typically larger than the FPS view. In other words, the viewers will less likely succeed in detecting the motion transitions as the size of the pedestrians decreases. Furthermore, the viewing angle of the strategy viewpoint (close to top view) makes the audiences more difficult to detect motion changes since smaller portion of the pedestrian body is visible. We also found an interaction between the two main effects ($F = 15.76, p < 0.018$) that decreasing gradient of the detected motion transitions in the strategy view is steeper than the FPS view (the top-left panel of Figure 2), especially, when the density is higher than 16. One possible explanation is from the fact that animated pedestrians heavily occlude each other in the FPS view when the density is high. The viewers tend to focus only on a small number of agents closer to the viewport even though the overall crowd density is increased. Based on this observation, 16 could serve as a good starting threshold for crowd scene using a FPS viewpoint.

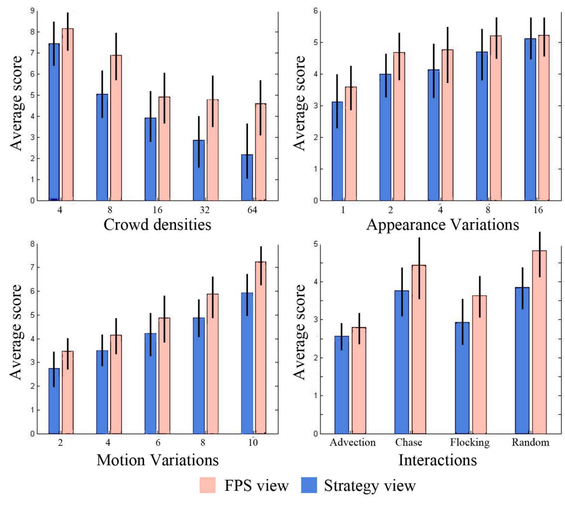

Figure 2: *Evaluation results: average scores & standard deviation errors of different crowd feature experiments*

2.3 Appearance Variety Effect

Diversifying both appearances and motions among crowd agents has significant influences on the animation realism of crowds. However, the interaction between the appearance and motion transitions is still an open problem. Previous work has showed that the appearance variation does not affect the detection of unchanged cloned walking motions [McDonnell et al. 2008]. This work focuses on the transitions when agents change their walking motions.

2.3.1 Experiment Description

A number of approaches have been proposed for generating appearance variations among crowd agents such as modifying agent colors and illuminations [Tecchia et al. 2002]. Since the motion capture data we used was captured from the same subject, we use the same 3D human model (about 800 triangles) with different texture maps to generate different appearances. A total of 16 distinct appearances are available in our psychophysical experiments (middle of Figure 1). When the lowest number of appearance is set to 1, that is, all the pedestrians look exactly the same (the second row of Table 2).

2.3.2 Results

The top-right panel of Figure 2 shows the average motion transition rates perceived by the participants. Similar to the crowd density experiment, the viewpoint was found to be the main effect ($F = 23.13, p < 0.008$), and the appearance number was another main effect ($F = 17.72, p < 0.014$). However, there was no evident interaction between appearance and viewpoint. Surprisingly, increasing the agent appearance number also increased the chance to detect unsmooth motion transitions even though the actual number of the motion transitions was kept unchanged. This might be due to some false detections of the motion transitions that was caused by visual interference from the increasing number of colors. Therefore, it is worthy to find out a good trade-off between the appearance variety and the smoothness of pedestrian motions in a simulated crowd.

2.4 Motion Variety Effect

The motion realism can be affected by the motion naturalness of each agent and motion variety among different agents. Pedestrians, in reality, may change their walking motions corresponding to different events, environments, and other pedestrians while keeping physically correct motion transitions. In computer simulation, these two factors have to trade off one another due to performance consideration. We evaluated the human perception of the motion changes when a certain number of motion candidates was given.

In this experiment, we tested different numbers of walking motion variety in a crowd. We hypothesized that, under the same transition frequency, audiences had a higher chance to detect motion transitions when a larger motion pool was used. For example, when comparing two crowds with 2 motions and 4 motions, respectively, we assume that a pedestrian performing a motion sequence “1234” has a better chance to be detected than another pedestrian performing a motion sequence “1212”. A total of 5 different sets of motion pools were used in this experiment (the third row of Table 2).

2.4.1 Results

The participants’ responses (the bottom-left panel of Figure 2) validated our assumption. While the viewpoint was still one of the main effects ($F = 27.08, p < 0.008$), the perception rates were more sensitive to the number of used walking motion candidates ($F = 37.76, p < 0.006$). However, we did not find significant interaction between the two. At both viewpoints, the perception rates were almost in direct proportion to the number of used walking motion candidates although the gradient under the FPS viewpoint was slightly higher. This result shows that increasing the motion variety of a crowd by switching among a larger number of motion candidates may amplify the impression of motion transitions of the entire crowd. Although we did not perform further experiments on the motion variety feature, one potential solution of this problem would be to employ more intelligent selection rules instead of random selections. Doing so, only transits among most similar motions will happen, which makes transitions as unnoticeable as possible.

2.5 Group Interaction Effect

In the previous sections, we evaluated the human perception of three quantitative parameters of pedestrian crowds. Another unique aspect of crowd animations, compared with single character animations, is collective interactions of the crowd. This qualitative feature represents the common purpose of the entire crowd or sub-groups in the crowd, and can induce strong attention from audiences. To this end, we simulated several crowd scenarios to evaluate the impact of such collective behavior on the detection of motion transitions. Specifically, we conducted user experiments to compare the random pedestrian scenario with the flocking, advection, and chase interaction scenarios (the fourth row of Table 2).

Figure 3: *The used group interaction scenarios: (Top-Left) Random, (Top-Right) Chase, (Bottom-Left) Flocking, and (Bottom-Right) Advection*

Random scene (top-left of Figure 3): It is a common scenario in crowd applications such as urban streets. Pedestrians move towards random destinations in a certain period through path navigation. All the experiments in the previous sections use random scenarios.

Chase scene (top-right of Figure 3): The viewpoint usually moves with the main character(s) while a group of supporting characters (enemies, followers, etc.) move around the main character(s). Audiences tend to give more attention on the main character(s).

Flocking scene (bottom-left of Figure 3): It is another widely used crowd simulation scenario such as emergency evacuation. Pedestrians in the crowd start from different locations and gather together to reach the same or similar purpose.

Advection scene (bottom-right of Figure 3): It can be described as a multiple group version of the flocking behavior and is often observed between two or more sub-groups in a crowd. Typical simulations include road intersections and battle fields. Pedestrians in a sub-group share the similar goal and head to other sub-groups.

2.5.1 Results

Based on the average rates over the different group interactions, it is clear that all three collective group behaviors, more or less, contribute to disguise motion transitions (refer to the bottom-right panel of Figure 2). The main effects were found to be group behavior type ($F = 44.56, p < 0.004$) and viewpoint ($F = 14.97, p < 0.012$). We did not evaluate the interaction between the group behaviors and viewpoints since the behavior types were not numerical variables. Among the three types of group interactions, **Chase** behavior showed a similar result with the **Random** crowd. The detection rate of motion transitions appeared slightly lower than that at the random crowd. Viewpoints (strategy or FPS) did not make significant differences in terms of the detection of motion transitions. One possible reason is that the leading character had a tendency to distract the viewers from detecting the motion transitions of other pedestrians, whereas the view camera moving with the crowd also made it easier to focus on a fixed number of pedestrians. **Flocking** behavior showed evident contributions to reduce the detection of motion transitions because of the high motion parallelism among pedestrians that were close to each other and heading towards the same destination. As a complex version of the flocking behavior, **Advection** interaction exhibited the highest impact on reducing transition detections at both viewpoints. One possible explanation is that the participants were highly distracted by the interleaving flow of the two sub-groups in the crowd, and thus had difficult time to focus on the motion changes of individual pedestrians.

3 Conclusions

In this paper, we evaluate the perceptual impact of motion transitions of individual pedestrians in a crowd. Unlike single character animations, the naturalness of motion transitions in a crowd animation can be affected by many aspects other than the transition algorithm itself. We performed a series of psychophysical experiments to investigate the influences of different viewpoints, crowd densities, appearance variations, motion variations, and collective group interactions. The main findings of our experiments can be summarized as follows:

1. Although we did not perform detailed quantitative experiments on many viewing angles and viewing distances as the work of [McDonnell et al. 2008], we found the strategy viewpoint is more effective to hide motion transitions than the FPS view in most experiments based on the average scores of evaluations (FPS scores are higher). It indicates that distant viewpoints are more effective in disguising unrealistic pedestrian motion transitions than close viewpoints.

2. Increasing the density of agent numbers in the viewport can significant help to hide motion transitions.

3. Adding more agent appearances (e.g., different textures) does not necessarily lead to better perception of motion transitions in a crowd. Conversely, it makes the motion transitions more noticeable to audiences, to a certain extent.

4. Increasing the number of motion candidates makes motion transitions among them easier to be detected even though the transition frequency remains unchanged.

5. The existence of collective behaviors or sub-group interactions can effectively decrease the negative impact of motion transitions among walking motions.

The above findings provide useful implications and practical guidelines to improve the perceptual realism of performance-oriented pedestrian crowds when inexpensive motion transition/blending algorithms are used.

4 Discussion and Future Work

In our study, we hypothesized that motion transitions are "undesirable" interpolations between good motions such as motion capture sequences. The transitions among walking motions are intentionally made noticeable to the participants when they are applied to individual agents. Although this allowed us to test the change of perceptions on a crowd, in some quality-oriented applications such as feature films, motion transitions are delicately computed based on sophisticated algorithms and are not necessarily worse than motion capture data. Increasing transitions may greatly enhance the overall crowd variety without noticeable jaggy motion effects. In this case, the transitions amplified by adding appearance and motion variations improve the overall visual realism of the crowd.

In the current study, we only tested interactions between viewpoints and the other three crowd features. An interaction was found between the viewpoint and the crowd density feature but not the other two. We plan to investigate the interactions among these three factors in future work. In addition, we are interested in probing the transition perceptions on other types of crowd motions in a more comprehensive study, such as running, talking, and fighting.

Acknowledgments

This work was supported in part by NSF IIS-0914965 and the Texas Advanced Research Program under grant No. 003652-0058-2007. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the funding agencies.

References

CUTTING, J. E., AND KOZLOWSKI, L. T. 1977. Recognizing friends by their walk: Gait perception without familiarity cues. *Bulletin of the Psychonomic Society 9*, 353–356.

KOVAR, L., GLEICHER, M., AND PIGHIN, F. 2002. Motion graphs. *ACM TOG 21*, 473–482.

MCDONNELL, R., LARKIN, M., DOBBYN, S., COLLINS, S., AND O'SULLIVAN, C. 2008. Clone attack! perception of crowd variety. In *SIGGRAPH '08*, 1–8.

MCDONNELL, R., LARKIN, M., HERNÁNDEZ, B., RUDOMÍN, I., AND O'SULLIVAN, C. 2009. Eye-catching crowds: saliency based selective variation. In *SIGGRAPH '09*, 1–10.

PELECHANO, N., ALLBECK, J. M., AND BADLER, N. I. 2007. Controlling individual agents in high-density crowd simulation. In *SCA '07*, 99–108.

REITSMA, P. S. A., AND POLLARD, N. S. 2003. Perceptual metrics for character animation: sensitivity to errors in ballistic motion. In *SIGGRAPH '03*, 537–542.

ROSE, C., GUENTER, B., BODENHEIMER, B., AND COHEN, M. F. 1996. Efficient generation of motion transitions using spacetime constraints. In *SIGGRAPH '96*, 147–154.

TECCHIA, F., LOSCOS, C., AND CHRYSANTHOU, Y. 2002. Visualizing crowds in real-time. *CGF 21*, 753–765.

VAN BASTEN, B. J. H., AND EGGES, A. 2009. Evaluating distance metrics for animation blending. In *FDG'09*, 199–206.