The Effect of Characters' Locomotion on Audience Perception of Crowd Animation

Wenyu Zhang(✉) and Nicoletta Adamo-Villani

Purdue University, West Lafayette, IN 47906, USA
zhan2586@purdue.edu

Abstract. A common practice in crowd animation is the use of human templates. A human template is a 3D character defined by its mesh, skeletal structure, materials, and textures. A crowd simulation is created by repeatedly instantiating a small set of human templates. For each instance, one texture is randomly chosen from the template's available texture set, and color and shape variety techniques are applied so that multiple instances of the same template appear different [1]. When dealing with very large crowds, it is inevitable to end up with instances that are exactly identical to other instances, as the number of different textures and shape modifications is limited. This poses a problem for crowd animation, as the viewers' perception of identical characters could significantly decrease the believability of the crowd simulation. A variety of factors could affect viewers' perception of identical characters, including crowd size, distance of the characters from the camera, background, movement, lighting conditions, etc. The study reported in this paper examines the extent to which the type of locomotion of the crowd characters affects the viewer's ability to perceive identical instances within a medium size crowd (20 characters). The experiment included 83 participants and compared the time participants took to spot identical characters in three different locomotion scenarios (e.g., standing, walking, and running). Findings show that the type of locomotion did not have a statistically significant effect on the time subjects took to identify identical characters within the crowd.

Keywords: Crowd animation · Virtual character · Perception

1 Introduction

Various crowd simulation techniques have been developed and are widely applied in the visual effects, animation, and video game industries. However, there is still a lack of research on the perceptual factors that could affect the audience experience of the crowd, such as the degree of realism of the characters, the level of detail, the crowd motions, etc. The work reported in the paper aims to fill this gap by examining the effects of characters' locomotion on the viewer's perception of identical characters in medium-sized crowd simulations.

M. Wölfel et al. (Eds.): ArtsIT 2021, LNICST 422, pp. 398–412, 2022.
https://doi.org/10.1007/978-3-030-95531-1_27

More specifically, this research focused on crowds which require a heterogeneous character appearance and motion. The goal of the study was to determine whether it is possible to use a lower number of entity characters depending on the particular type of locomotion performed by the crowd. The findings from the study have practical implications for real-time simulations (e.g., computer games) as well as off-line rendering simulations (e.g., crowds in film). One of the main goals in animation production is to maximize the visual quality of moving pictures with a minimal expenditure of technical resources. The results from this study can help achieve this goal. They can help those involved in animation production save time and resources by decreasing the number of different human templates that are necessary in certain crowd simulations.

2 Literature Review

2.1 Character Animation

The production process of character animation can be broken down into five basic parts: modeling, texturing, rigging, animating and rendering. Modeling, texturing, and rendering all determine the physical appearance of characters, while rigging and animating control the characters' movements and facial expressions.

Modeling is a process whereby the creator defines the shape of the characters without giving consideration to their texture. It allows the creator to display several basic properties of a character, including height, gender, age, body shape, hair style, and muscle level. Zell *et al.* [2] noted that "shape is the main descriptor for realism, and material increases realism only in case of realistic shapes."

Rigging and Animating are two key factors in crowd animation. Traditionally speaking, the character models in crowd animation are polygonal meshes rigged by bones. When the joints are rotated, the vertices cluster attached to the joints becomes deformed along a predetermined trajectory, which is how character animation is generated [3]. In real world production, character motion can be created either by animators' key-framing and adding frame interpolation, by using physics-based animation generated by computer simulation tools, by capturing real-time motion data from devices on actors (motion capture), or via any combination of the three aforementioned techniques [4]. However, certain comprehensive methods need to be adopted in creating crowd animation because moving crowds involve complicated mechanics which require algorithmics [5]. Such methods, which go beyond an individual character's locomotion, have been studied by previous researchers. For example, walking is a common mode of locomotion that can easily be produced for an individual character.

However, in the case of a group of walking characters (e.g., pedestrians on the street), factors such as collision avoidance must also be considered [6]. Additionally, many algorithms related to the motion trajectories of crowds have been developed in recent years. For instance, Yu and Terzopoulos [7] developed a novel framework for pedestrian characters, including behavioral interaction in urban settings. Guy *et al.* [8] presented a technique called Personality Trait Theory to create heterogeneous crowd motion. Sun *et al.* [9] simulated realistic crowd trajectory in an urban scenario surrounded by traffic, vehicles, intersection, etc.

In a crowd simulation, characters' locomotion and behavior inevitably rely on the nature and quality of the algorithms operating behind the scenes.

2.2 Crowd Rendering

3D renderers are tools used to output final image sequences in animation production. They generate various results depending on the shading models and numerous physical parameters inputted by the animators. Traditionally, crowd animation was rendered through animation programs. However, with the development of 3D games, many programs tend to have the capability to simulate crowds in real time. In the earlier studies on crowd rendering, limited computational speed was the main problem in the creation of 3D scenes with populated characters. Many acceleration techniques for the rendering of large environments were subsequently invented. A technique called "instance" has been frequently used in crowd simulation. In a shading API (Application Programming Interface) such as OpenGL or DirectX, a geometry shader can be used to deform the vertices and the triangle mesh of a crowd with only one call in GPU (Graphics Processing Unit) [10]. Ashraf and Zhou [11] applied a hardware-accelerated method through programmable shaders to animated crowds. In Peng *et al.* [12], developments have been made to utilize GPU in their parallel architecture to improve the performance of graphics computation. They invented a mesh simplification algorithm which can render a real-time crowd system on the GPU. Klein *et al.* [13] created an innovative method which allows instances of 3D characters with controllable parameters to be rendered on the web. In the work of [3], a novel crowd rendering system that simultaneously runs real-time on GPU and decreases the computation load on the graphics card was developed. The scene includes 30,000 instances in real-time motion.

Shading is also a significant component in rendering real-time images. Maciel and Shirley [14] implemented the LOD (Level-of-Detail) technique to create impostors to reduce the complexity of rendering. This method later evolved to an IBR (Image-based Rendering) technique which was adopted by many researchers. In a study by Tecchia and Chrysanthou [15], the Image-based Rendering (IBR) method was adopted, and the characters were pre-computed and animated. A multi-pass algorithm was first used to retouch different parts on the character, followed by the addition of efficient shading and shadow. In a study by Tecchia *et al.* [16], they used an approach whereby each character was transformed into an image-based impostor which possesses an adaptive resolution depending on one's viewing angle. Ciechomski *et al.* [17] presented a customized hardware rendering pipeline which created texture variety from a single texture in HSB color space. Millan and Rudomin [18] combined the imposter and instancing techniques and created a program which is more efficient in rendering large crowds.

2.3 Perception of Virtual Characters

Ciechomski *et al.* [17] has stated that "For a human crowd, variation can come from the following aspects: gender, age, morphology, head, kind of clothes, color of clothes and behaviors." In other words, the perception of human crowds in animation mainly depends on two aspects - appearance and behavior.

All the virtual CG characters can be classified into two categories: photo-realistic and stylized. In a study by Zell *et al.* [4], it was found that factors such the shape of a character's body and its material (especially the albedo texture) can significantly affect audience perception. These two factors have a strong influence on how realistic the characters are perceived to be.

Another factor that affects the believability of perception is the facial proportion of characters. Green *et al.* [19] concluded that facial height, jaw width, and eye separation are all considered to be important factors which can increase the appeal of animated characters.

Besides their exterior appearance, the behavior or motion of characters likewise plays an important role in creating realistic perceptions. Based on a study [20], when the characters are in motion (e.g., walking or running) as opposed to staying still, viewers can appreciate that the virtual characters resemble real world human beings, instead of perceiving them as a group of static dot-shape objects. Research by McDonnell *et al.* [21] compared the reaction times in spotting appearance-based duplicated characters versus motion-based duplicated characters. They concluded that characters cloned by appearance are more conspicuous than characters cloned by motion. Also, they discovered that the position layout of characters affected the viewers' perception - horizontal layout makes it easier for the audience to spot cloned characters compared to a vertical or diagonal layout. One limitation of their experiment is that all the testing characters were positioned facing forward, which is not considered typical in crowd animations. Pražák and O'Sullivan [22] studied the locomotion variety in crowd animation perception. They adopted motion capture techniques to capture 83 actors' real-world motion data (including both males and females) and created a virtual scene to perform the experiment. They claimed that at least three different locomotion types are needed to be displayed for each gender to achieve a realistic level of behavioral variety in a pedestrian scene. However, their character set was relatively small, with only 24 characters being shown at a time in each scene. Moreover, they did not examine the effects of the various types of motion in the experiment.

Eye tracking has become quite popular in perception studies in recent years. Using an eye tracking device, McDonnell *et al.* [23] found that head and upper body are the first part viewers tend to notice, regardless of the character's position, motion, gender, size, etc. They also found that creating more kinds of head accessories and variable top textures is more effective at increasing variety than alternating the facial geometry of characters.

When it comes to facial close-ups, the eyes tend to catch viewers' attention more than other body parts. A recent study [24] confirmed that viewers primarily maintain their glance at the virtual characters' eyes and mouth. On average, it was found that participants spend around 35% of the time looking at the eyes, while spending no more than 10% of the time focusing on other parts of the body.

Figure 1 is a screenshot of an animated commercial short for Westfield Stirling Shopping Mall [25]. Some of the CG characters are walking randomly in the mall; while some are standing still. They all have different appearance and slightly difference behavior which increases the perception fidelity of crowd animation.

Fig. 1. A screenshot of the visualization project Westfield Stirling by New Holland Creative.

3 Methodology

The goal of this study was to determine whether different types of locomotion would affect viewers' perception of the crowd. The participants watched randomized video clips representing three scenarios and were then instructed to complete a related online survey. The study adopted a quantitative research approach that compared the length of time that participants spent on each scenario to identify identical characters. A customized Bayesian Linear Mixed Model was employed to analyze the collected data.

The independent variables in this research were the type of locomotion (standing, walking, running) of the 3D characters in the crowd and the gender of the participants. The dependent variable was the length of time subjects took to identify two identical characters in the crowd.

3.1 Hypotheses

H_{01}: Participants will spend the same amount of time to identify identical characters in all the three locomotion scenarios.
H_{a1}: Participants will spend different amount of time to identify identical characters in each of the three scenarios. Specifically, participants will spend more time to identify identical characters in the Running Scenario than in the Walking and Standing Scenarios, respectively.
H_{02}: Participants will spend the same amount of time to identify identical characters regardless of the participants' gender.
H_{a2}: The time participants will spend to identify identical characters will vary depending on the participants' gender.

3.2 Subjects

A total of 83 participants took part in this study. Thirty-three participants were students from the Computer Graphics Technology department at Purdue University. Fifty participants were selected via a survey posted on Amazon Turk. The participants were recruited without regard to gender and resulting in 46 males and 37 females in the pool. Participants' age ranged from 18 to 64 years old. Participants' familiarity with computer animation ranged from zero experience to very familiar with computer animation. All the participants could see the computer screen clearly, with or without corrective lenses.

3.3 Stimuli

The stimuli used in this study consist of three online videos demonstrating different types of character locomotion within crowd animation, along with an online survey. The crowd animation video clips were created using Maya 2016 with Golaem plugin and were rendered using Mental Ray renderer. The rendered videos contain both highlight and shadow in order to simulate realistic lighting. However, the materials on the characters do not include any other channels besides diffuse textures. All the characters' exterior, such as garment texture, is from the preset package of Golaem plugin. The characters' locomotion (e.g., walking, running) was also created using Golaem presets. The characters' moving trajectories are customized to allow the characters to have specific paths without moving out of the frame. Also, to assure all the other parameters stayed uniform, the camera angle, lighting, shadow, contrast, are set up completely identical in each video clip. The camera is positioned at one side of the scene with a tilting angle of 30° towards the ground. The lens has a view angle of 35° to capture the full scene.

In each scene, there are 18 characters with heterogeneous appearance and only two characters with homogeneous appearance, which includes skin color, hair color, color of shirt, pants and shoes. In the standing scenario, characters stand still on the ground surface and exhibit casual turning-in-place movements.

In the walking scenario, characters walk in random trajectories on the ground surface. In the running scenario, characters run around in random trajectories.

After the animation was completed, all frames were exported from the animation package. Image sequences for each scenario were processed in video editing programs and output as three 98-s video clips. Each video clip had a 10-s opener with instruction reminding viewers to be prepared for the experiment. Each video clip looped 10 times itself. All the experiment videos were in sRGB color space without any post-processing or visual effects.

The three video clips for experiment were uploaded to online video platform Vimeo which can be viewed through following link. Screenshots from the three videos are shown in Fig. 2.

- https://vimeo.com/367614577
- https://vimeo.com/367615337
- https://vimeo.com/367613352

Fig. 2. Locomotion scenarios illustrating standing, walking, running, respectively.

3.4 Evaluation Instrument

This experiment required participants to view a series of animation video clips. Therefore, a laptop or personal computer with proper display and fast Internet access was required. Mobile devices were not allowed in this study given that the screen resolution on such devices and the various nuances of person-device interaction might affect the perception results.

Data collection was performed via an online survey created using the Qualtrics survey platform. The videos were embedded into the survey platform and all user interaction controls were disabled. The survey included the IRB consent form, detailed experimental instructions, a demo video, three formal testing videos, a demographics questionnaire, and optional feedback.

Since this study required quick responses from the participants, detailed instructions along with a video tutorial were displayed to participants at the beginning of the experiment to ensure they thoroughly understood the experimental procedure. In addition, a demo video was presented to allow participants to familiarize themselves with the procedure and promote reliable results. Participants were expected to adjust page zoom to a suitable resolution in the browser to allow them to watch the entire frame.

Formal video clips for testing began to play automatically as soon as the participant displayed the page. Each video clip had a text reminder stating: "Please move the cursor on the blue button. (Do not click until you have found two identical characters)." Along with each formal experiment video, there were required questions on the following page letting participants select the identical character, if found. Each question had only one correct answer out of three choices. The answer did not contain any text but only a pair of screenshots of the characters (full body front and back) appeared in the video. Thus, viewers might have had a more intuitive impression to select the character they believed they have found. Participants were forced to select an answer before they could jump to the next page.

In order to decrease potential confounds stemming from the learning effect (whereby participants' performance improves over time as they are exposed to the same stimulus), the order of the three video scenarios was randomized. We randomized the video groups into three different combinations to make sure each scenario would not always appear at the first. This greatly reduced the audience's learning effect. The order combinations were Standing-Walking-Running, Walking- Running-Standing and Running-Standing-Walking.

3.5 Procedure

For each scenario, the video clip started to play automatically and looped for 15 times. All the interaction controls were disabled on the videos. Thus, participants were not able to pause, adjust speed, download, or loop the video by themselves. Participants were asked to click on the blue button showing "CLICK ME" at the bottom right corner of each scenario page as soon as they spotted the two identical characters. The system recorded the exact response time for each participant. Figure 3 shows a screenshot of the Walking Scenario stimuli.

Next, participants were asked to select which of the three types of characters were identical in the video clip. After a selection was made, the page would progress to the next video. After viewing all the video clips and answering the pertaining questions, the participants were asked to fill out a brief demographic questionnaire. It collected participants gender, age and their familiarity of computer animation. Finally, they were given the option to share any feedback or comments they may have had regarding their experience before concluding the study.

4 Data Analysis

After the experiment was conducted, participant response times (i.e., the amount of time each participant spent to identify identical characters in each video) were collected. Since there were fixed and multiple random factors in this study, a Bayesian Linear Mixed Model was used to determine whether the response times varied significantly across the three locomotion scenarios (standing, walking, and running).

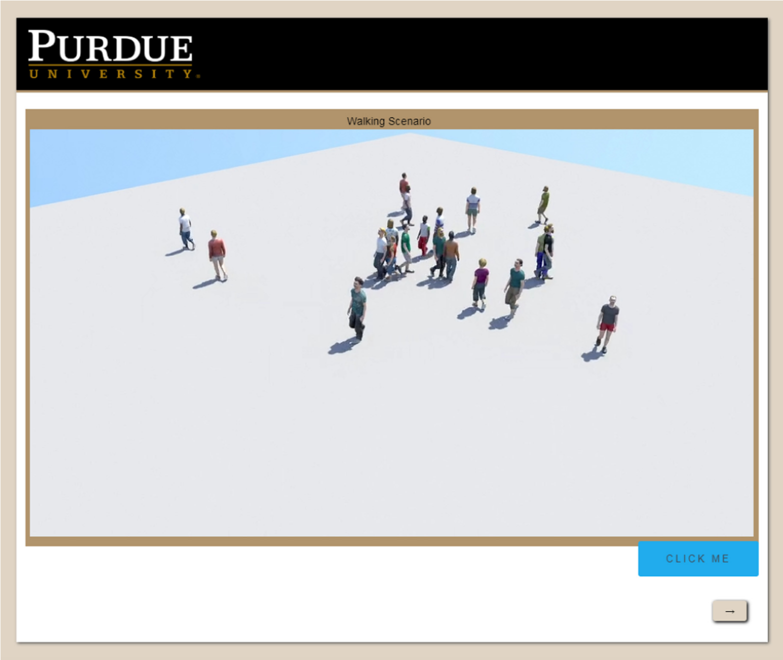

Fig. 3. A screenshot of formal testing video clip (walking scenario).

4.1 Data Pre-processing

The dependent variable in this study was the response time, or the length of time that participants spent on each scenario before clicking the mouse (indicating that they identified two identical characters). First, an accuracy check was performed to clean up the collected data; participants who selected incorrect answers were subsequently removed from data set. Standing Scenario had an accuracy of 75%; Walking Scenario had lowest accuracy of 62%; Running Scenario had highest accuracy of 87%. Figure 4 is a bar graph to visualize the accuracy result. Since the actual video would not play until the 10^{th} second and would terminate at the 98^{th} second, participants' who had spent less than 10 s and greater than 98 s in watching each video clip were removed from data set. After the clean-up, there were 51 available responses in the data set, 28 from males and 23 from females. The reaction times across the three video types were then analyzed using Bayesian Linear Mixed Model.

4.2 Data Model

Each participant in our study was exposed to all three video categories. Only the subjects who identified every pair correctly and responded within the acceptable range of response times (as explained above) were included in the response time

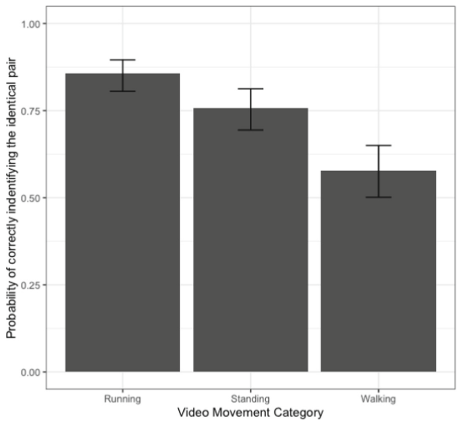

Fig. 4. Response accuracy of each locomotion scenario.

analysis. Combining all the factors which might affect the result of this study, we attempted to fit the model as below:

$$Time_{ijk} = \mu + Video_i + Subject_k + Period_j + Sequence + Gender + \epsilon_{ijk}$$

where:

1. $Time_{ijk}$ is the actual response time for subject k watching video i in time period j.
2. μ is the overall mean expected response time.
3. $Video_i$ is the effect of the i^{th} video category (Running, Walking, Standing) on the expected response time.
4. $Subject_k \sim N(0, \sigma^2_{subj})$ is the random effect of subject k on expected response time.
5. $Period_j$ is the effect of the j^{th} time period on the expected response time.
6. $Sequence$ is the effect of video display order on the expected response time.
7. $Gender$ is effect of different gender on the expected response time.
8. $\epsilon_{ijk} \sim N(0, \sigma^2)$ is the error between expected and actual response time.

The model includes fixed-effects stemming from our independent variable video category (standing, walking, running) and factors corresponding to video order, period, gender as explained above. In addition, we included random subject effect, in order to control for heterogeneity of each subject.

In this study, there were three categories under $Video_i$, 51 different individuals under $Subject_k$, three categories under $Period_j$, three different orders under $Sequence$, and two categories under $Gender$. $Video_i$ includes VideoS, VideoW, VideoR; $Period_j$ includes Period1, Period2, Period3; $Sequence$ includes Sequence1, Sequence2, Sequence3; $Gender$ includes GenderMale and GenderFemale.

4.3 Data Analysis

As the graph suggests, participant response accuracy was highest in the running category, followed by standing and walking scenarios, respectively.

Using Bayesian Linear Mixed Model, such result was yielded from the data model with a 95% credible interval ranging from 2.5% to 97.5% (Table 1).

Table 1. Multiple-factor credible interval

	.lower	.upper	.width	.point	.interval
VideoS	−18.0	13.9	0.95	median	qi
VideoW	−2.75	28.1	0.95	median	qi
Period2	−22.3	9.85	0.95	median	qi
Period3	−24.3	7.85	0.95	median	qi
Sequence2	−23.6	12.7	0.95	median	qi
Sequence3	−22.5	15.3	0.95	median	qi
GenderMale	−31.1	−0.44	0.95	median	qi

In this case, VideoR, Period1, Sequence1 and GenderFemale are used as baselines. The most plausible values with higher probability of representing the true estimate indicate that the mean of the intervention group VideoS and VideoW should be either lower or higher compared to the comparison group VideoR. As 0 lies within the interval, we do not have statistically significant evidence to claim that there is difference between VideoS, VideoW, and VideoR.

Credible interval for Period2 and Period3 contains 0. This indicates we do not have statistically significant evidence to claim that there is difference between Period1, Period2 and Period3. Accordingly, credible interval for Sequence2 and Sequence3 contains 0. This indicates we do not have statistically significant evidence to claim that there is difference between Sequence1, Sequence2, and Sequence3, either.

However, GenderMale has both negative lower bound and upper bound which does not contain 0. Thus, *Gender* turned out to be significant factor in this data model. Two interaction plots regarding $Video_i$ and *Gender* were generated after this interesting finding as shown in Fig. 5. In the first plot, it shows that male participants always had shorter response time than female participants across all the three video types, especially in Standing and Running Scenario. In the second plot, female participants tended to have lower variance while male participants had a higher variance. However, both genders performed worst in Walking Scenario.

4.4 Results

Results from the data analysis showed that the time participants took to identify two identical characters in the crowd were not significantly affected by different locomotion categories. Hence, we failed to reject the null hypothesis. There

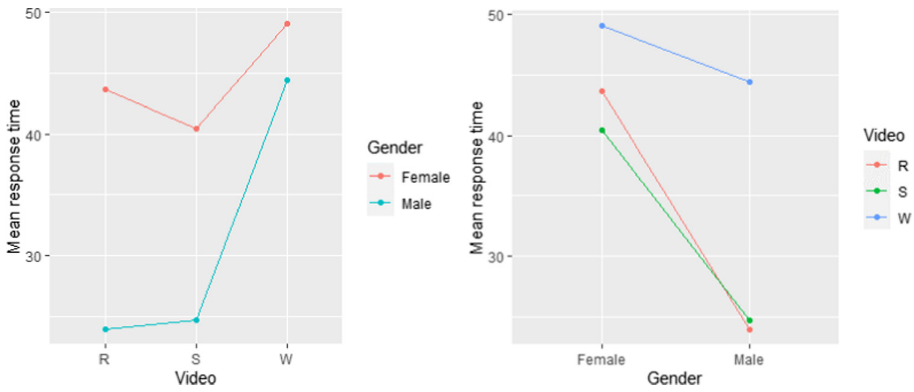

Fig. 5. Visualization of response time per gender.

was no significant difference in reaction time across the three different crowd animation scenarios. However, gender had a significant effect on participants' perception of identical characters within the crowd. Male viewers tended to be able to spot identical characters quicker than female viewers. In the three types of scenarios, male and female viewers had smaller difference in Walking Scenario while they had major difference in Standing and Running Scenario.

5 Discussion and Future Work

The results of the experiment reported in the paper indicate that the type of characters' locomotion (e.g., standing, walking, or running) in a medium size crowd animation has no significant effect on the audience's perception of the crowd. In particular, the type of locomotion exhibited by the characters in the crowd scenario does not significantly impact the time people take to spot identical characters. Findings also suggest that the gender of the participants has an impact on perception of crowd animation, with male participants being able to spot identical characters in a crowd more quickly than female participants.

The findings of the study have important practical implications for animation production. They can help animators save time and resources by optimizing the number of different human templates that are necessary in certain scenarios of crowd simulation (e.g., stadium audience, city street pedestrians, architecture visualization, etc.).

This study had several limitations and potential confounds which could be overcome in future experiments. First, a power analysis was not performed prior to the actual experiment. The researcher used as much of the subject's background characteristics and demographics in the design and analysis of the study to obtain as much power as possible under the circumstances. The pool of subjects was fairly representative of the target population, as it included participants of different ages ranging from 18 to 54 years old and with a wide range of animation experience.

Second, the position of the characters in the crowd at any given moment of time might have had an effect on participants' perception. For example, identifying two identical characters that happened to be running close to each other may have been easier than if the characters were far apart. Thus, distance between two identical characters could have been a significant factor that affected the perception in such scenarios.

Third, all the shots were static without any camera movement, which is not always true in real world films. In a case with camera movement (e.g., a top-down view with a dolly shot), the audience might not be able to focus on a specific area. Hence, the probability that viewers spot identical characters may be lower.

Fourth, the videos used in this experiment were quite rudimentary and considerably lower in quality compared to real-world commercial film productions. Visual fidelity was relatively low due to quality of character texture assets and lack of surrounding environment. The videos also lacked elements used in compositing such as smoke, fog, haze, dust, and flares - all of which are inevitably present in the real world. Further, all the testing scenarios did not include any 3D objects which might become blockers (e.g., buildings, poles, signs), but only an open space on a flat ground. As a result, the audience might be able to perceive identical characters more quickly and easily in our study as compared to real-world animated films.

Fifth, a phenomenon known as the learning effect might have also played a role in this experiment. Participants might have been able to achieve better results with more and more familiarity with the testing procedure in a short period of time. The researcher used randomization to mitigate this effect. A demo video was given at the beginning of the study, so participants could become familiar with spotting identical characters before conducting the actual experiment.

Finally, viewers' perception of the characters might have been affected by the intrinsic design features of the characters, in addition to our variable of interest (locomotion). For example, it is known that human eyes are more sensitive to certain colors of the visible spectrum (e.g., solid red and yellow) than to others, and so participants' response times might have been affected by the different colors of the characters.

In future experiments, characters' motion paths could be varied to exhibit different trajectories. For example, all the characters could be running towards the same target, or all of them could be running around in a loop. It would be interesting to see whether the moving path of the crowd as a whole would affect viewers' perception of identical characters.

In addition, certain camera angles, such as the absolute top view, could make it very difficult to spot identical characters. The difficulty of perception would also depend on the distance between the rendering camera and the characters. Further, it would be worthwhile conducting research on crowd perception under moving cameras.

Future experiments could also diversify characters' appearance, so that differences in skin color, gender, body shape, and other variables can be included and their effects on audience perception could be analyzed. Characters could also be made to wear glasses, hats, and other accessories to investigate their effects on viewers' perception.

References

1. Thalmann, D., Musse, S.R.: Crowd Simulation, 2nd edn. Springer, London (2013). https://doi.org/10.1007/978-1-84628-825-8
2. Zell, E., Zibrek, K., McDonnell, R.: Perception of virtual characters. In: SIGGRAPH 2019: ACM SIGGRAPH 2019 Courses, vol. 21, pp. 1–17 (2019). https://doi.org/10.1145/3305366.3328101
3. Dong, Y., Peng, C.: Real-time large crowd rendering with efficient character and instance management on GPU. Int. J. Comput. Games Technol. **2019**, 1792304 (2019). https://doi.org/10.1155/2019/1792304
4. Zell, E., et al.: To stylize or not to stylize? The effect of shape and material stylization on the perception of computer-generated faces. ACM Trans. Graph. **34**, 1–12 (2015). https://doi.org/10.1145/2816795.2818126
5. Lemercier, S., et al.: Realistic following behaviors for crowd simulation. Eurographics **31**, 489–498 (2012). https://doi.org/10.1111/j.1467-8659.2012.03028.x
6. Reynolds, C.W.: Flocks, herds, and schools: a distributed behavioral model. Comput. Graph. **21**(4), 25–34 (1987). https://doi.org/10.1145/37402.37406
7. Yu, Q., Terzopoulos, D.: A decision network framework for the behavioral animation of virtual humans. In: Metaxas, D., Popovic, J. (eds.) Eurographics/ACM SIGGRAPH Symposium on Computer Animation, pp. 119–128 (2007). https://doi.org/10.5555/1272690.1272707
8. Guy, S.J., Kim, S., Lin, M.C., Manocha, D.: Simulating heterogeneous crowd behaviors using personality trait theory. In: Bargteil, A., Panne, M. (eds.) Eurographics/ACM SIGGRAPH Symposium on Computer Animation, pp. 43–52 (2011). https://doi.org/10.1145/2019406.2019413
9. Sun, L., Li, X., Qin, W.: Simulating realistic crowd based on agent trajectories. Comput. Anim. Virtual Worlds **24**, 165–172 (2013). https://doi.org/10.1002/cav.1507
10. Carucci, F.: GPU Gems 2, pp. 47–67. Addison-Wesley, Boston (2005)
11. Ashraf, G., Zhou, J.: Hardware accelerated skin deformation for animated crowds. In: Cham, T.-J., Cai, J., Dorai, C., Rajan, D., Chua, T.-S., Chia, L.-T. (eds.) MMM 2007. LNCS, vol. 4352, pp. 226–237. Springer, Heidelberg (2006). https://doi.org/10.1007/978-3-540-69429-8_23
12. Peng, C., Park, S.I., Cao, Y., Tian, J.: A real-time system for crowd rendering: parallel LOD and texture-preserving approach on GPU. In: Allbeck, J.M., Faloutsos, P. (eds.) MIG 2011. LNCS, vol. 7060, pp. 27–38. Springer, Heidelberg (2011). https://doi.org/10.1007/978-3-642-25090-3_3
13. Klein, F., Spieldenner, T., Sons, K., Slusallek, P.: Configurable instances of 3D models for declarative 3D in the web. In: Proceedings of 19th International ACM Conference on 3D Web Technologies, Vancouver, pp. 71–79 (2014). https://doi.org/10.1145/2628588.2628594
14. Maciel, P.W.C., Shirley, P.: Visual navigation of large environment using textured clusters. In: 1995 Symposium on Interactive 3D Graphics, pp. 95-ff (1995). https://doi.org/10.1145/199404.199420

15. Tecchia, F., Chrysanthou, Y.: Real-time rendering of densely populated urban environments. In: Péroche, B., Rushmeier, H. (eds.) EGSR 2000. Eurographics, pp. 83–88. Springer, Vienna (2000). https://doi.org/10.1007/978-3-7091-6303-0_8
16. Tecchia, F., Loscos, C., Chrysanthou, Y.: Image-based crowd rendering. IEEE Comput. Graph. Appl. **22**, 36–43 (2002). https://doi.org/10.1109/38.988745
17. Ciechomski, P.H., Schertenleib, S., Maïm, J., Maupu, D., Thalmann, D.: Real-time shader rendering for crowds in virtual heritage. In: Mudge, M., Ryan, R.N., Scopigno, R. (eds.) The 6th International Symposium on Virtual Reality, Archaeology and Cultural Heritage, pp. 1–8 (2005). https://doi.org/10.2312/VAST/VAST05/091-098
18. Millan, E., Rudomin, I.: Impostors and pseudo-instancing for GPU crowd rendering. In: Proceedings of 4th International Conference on Computer Graphics and Interactive Techniques in Australasia and Southeast Asia, Kuala Lumpur, pp. 49–55 (2006). https://doi.org/10.1145/1174429.1174436
19. Green, R.D., MacDorman, K.F., Ho, C., Vasudevan, S.: Sensitivity to the proportions of faces that vary in human likeness. Comput. Hum. Behav. **24**, 2456–2474 (2008). https://doi.org/10.1016/j.chb.2008.02.019
20. Johansson, G.: Visual perception of biological motion and a model for its analysis. Percept. Psychophys. **14**, 201–211 (1973). https://doi.org/10.3758/BF03212378
21. McDonnell, R., Larkin, M., Dobbyn, S., Collins, S., O'Sullivan, C.: Clone attack! Perception of crowd variety. ACM Trans. Graph. **27**, 1–8 (2008). https://doi.org/10.1145/1360612.1360625
22. Pražák, M., O'Sullivan, C.: Perceiving human motion variety. In: Proceedings of ACM SIGGRAPH Symposium on Applied Perception in Graphics and Visualization, pp. 87–92 (2011). https://doi.org/10.1145/2077451.2077468
23. McDonnell, R., Larkin, M., Hernández, B., Rudomin, I., O'Sullivan, C.: Eye-catching crowds: saliency based selective variation. ACM Trans. Graph. **28**, 1–10 (2009). https://doi.org/10.1145/1531326.1531361
24. Schwind, V., Jäger, S.: The uncanny valley and the importance of eye contact. In: Mensch und Computer 2015 - Tagungsband, pp. 153–162 (2015). https://doi.org/10.1515/9783110443929-017
25. Westfield Stirling short film. https://vimeo.com/317897403