Perceived realism of pedestrian crowds trajectories in VR

ANONYMOUS AUTHOR(S)

SUBMISSION ID: 5815*

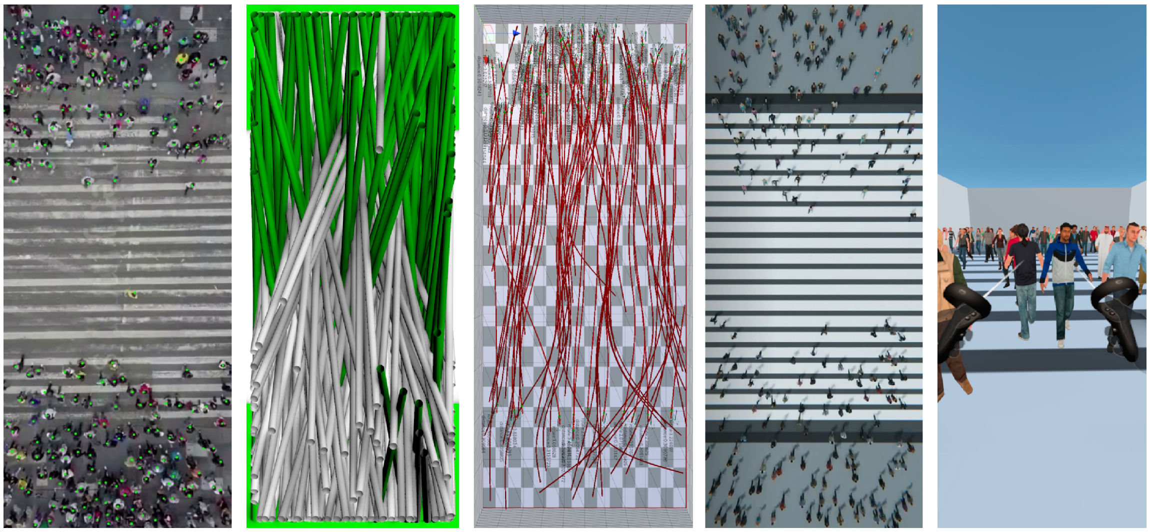


Fig. 1. From left to right, tracking trajectories from real data, rectilinear trajectory generation, animation process for real trajectories, aerial view simulation, simulation in immersive environment.

Crowd simulation algorithms play an essential role in populating Virtual Reality (VR) environments with multiple autonomous humanoid agents. The generation of plausible trajectories can be a significant computational cost for real-time graphics engines, especially in untethered and mobile devices such as portable VR devices. Previous research explores the plausibility and realism of crowd simulations on desktop computers but fails to account for the effect of immersion. Unlike desktop simulations, users of VR simulations can use stereo and motion parallax to make judgements about plausibility of the environment. This study explores how immersion in VR affects the perceived realism of crowd trajectories. We do so by running a psychophysical experiment in which participants rate the realism of real/synthetic trajectories data. Results show that trajectories from real data and synthetic rectilinear constant speed trajectories achieve similar levels of perceived realism.

CCS Concepts: • **Computing methodologies** → **Agent / discrete models**; • **Human-centered computing** → **User studies**.

Additional Key Words and Phrases: crowd simulation, perception, virtual reality

ACM Reference Format:

Anonymous Author(s). 2018. Perceived realism of pedestrian crowds trajectories in VR. In *Woodstock '18: ACM Symposium on Neural Gaze Detection, June 03–05, 2018, Woodstock, NY*. ACM, New York, NY, USA, 15 pages. https://doi.org/10.1145/1122445.1122456

1 INTRODUCTION

Crowd simulation is the complex process of simulating collective movement of several individual entities such as virtual humans or animals. It is used in numerous contexts, such as urban or evacuation planning, video games, and cinematography. For realistic, high-density crowd simulations the computation and memory cost is significant and requires the use of high-end desktop computers [7, 33] or even in some cases distributed computing [17, 34]. With Virtual Reality (VR) becoming an increasingly popular technology, crowd simulation algorithms play an essential role in enabling the creation of rich VR experiences. In addition, the recent emergence of portable VR devices with limited computational resources, such as Oculus Quest and Quest 2, challenge developers as their CPU and GPU capabilities are significantly lower than current desktop-class CPU and GPUs. How does a user perceive a crowd simulation as plausible or realistic? This question is challenging as multiple simulation stages are involved, such as rendering, animation, pedestrian dynamics and, sometimes, the observer's interactions with the crowd such as dynamic avoidance of the observer. In particular, crowd simulation algorithms deal with different aspects whose realism can be evaluated either jointly or individually: high-level planning, global path planning, path following, local movement, locomotion synthesis and animation. Evaluating the realism of a crowd is a challenging task, and many attempts were made over the years. Prior studies introduce metrics [11, 28], rely on subjective evaluation by performing user tests[5, 25] or even compare against real-world crowd data [4, 8, 16]. Some of them vary parameters such as camera position and orientation [1, 5], environments [22], density levels of the crowd [15, 19], situations or contexts (for example an evacuation process or a simple walk in the park) [27], and algorithms [12]. While previous studies explore the evaluation of perceptual realism of crowd simulations using desktop computers, or in some cases VR headsets, to the best of our knowledge there is no study that addresses *perceptual realism of crowd simulation trajectories in large and high-density crowds from an immersion perspective*. Our paper is motivated by the fact that the computation of realistic trajectories in a densely crowded environment might have high computational cost that could be prohibitive to use on untethered VR devices. Determining if a computationally cheap algorithm can create realistic trajectories when the user is immersed in a crowded context is crucial since it saves CPU and/or GPU time. The contribution of our study is two-fold: firstly, we identify and validate through a user study (Experiment 1) this computationally inexpensive method. Validation consists of measuring perceived realism to assess the generation of a plausible set of trajectories for a high-density pedestrian crossing scene. The first experiment consists of watching crowd simulation videos from different points of view, following by a perceived realism questionnaire. Secondly, we compare via a VR user test (Experiment 2) the perception of realism between the simulated trajectories and a set of real-world trajectories traced from video footage. The second experiment measures perceived realism from an immersive viewpoint by using the same questionnaire of Experiment 1 plus common HCI questionnaires about immersion [31].In addition, we evaluate the crowd density perception with the participant immersed in such environment, both for real and simulated data. For Experiment 2, we designed a full-VR experience where users can observe crowds driven by real trajectories or rectilinear trajectories (as shown in Fig. 1). Both experiments are implemented so that they can be run remotely without supervision.

2 LITERATURE REVIEW

The crowd simulation community has used several approaches to assess the quality of crowd simulations [24]. Such approaches include customized metrics that are meaningful to the task [11, 26, 28], user studies in various contexts [3, 25] and comparisons of simulations with real-world data [4, 8, 16, 36]. In the following, we mostly focus our discussion on works that study factors that affect the perceptual realism of crowds in VR, as this is the most relevant to the work presented in this paper.

One strand of prior work explores the perceptual realism of crowds via psychophysical experiments concerning different aspects such as social forces, social interactions and group dynamics. For example, Ennis *et al.* [5] explore how scene context and viewpoint affect perceptual realism of static crowds. In particular, this work uses rendered images from various camera positions and orientations and measure the impact of plausibility for static crowds. O'Connor *et al.* [20] observe how social forces affect the perceptual realism of crowds trajectories from an aerial view while Hoyet *et al.* [10] study the impact of shoulder motion on perceptual realism and show that shoulder movements enhance perceived realism of dense crowds. Kyriakou *et al.* [15] show how collision avoidance and other social interactions such as gaze and salutations affect the perceived plausibility of a crowd in a VR setting. Barut *et al.* [2] examine the elicitation of manoeuvring motion illusion from rectilinear trajectories [14]. Such illusion occurs when rectilinear trajectories are perceived as wriggling trajectories. In other words, the observer perceives variations in the linear trajectories that are not present. In the context of crowd simulations, this effect increases realism because pedestrian trajectories are generally not linear or constant speed. O'Connor *et al.* [22] explore how agent grouping dynamics affect the perceptual realism crowd simulations by varying crowd density and frequency in different urban locations.

2.1 Measuring Realism of Crowd Trajectories

There are several approaches to measure the similarity between simulated data and real-world data. Guy *et al.* [8] for example define an entropy metric that compares entire simulations to reference data. Charalambous *et al.* [4] use the principle of Pareto optimality to compare between crowd data under multiple (often conflicting) criteria, whereas Wang *et al.* [35] define an inference-based similarity metric. Wolinski *et al.* [36] propose a genetic algorithm to find optimal parameters for a crowd simulator so that simulations match specific features of reference data; this in turn allows for fairer comparison between simulators. Conversely, several studies are based on psychophysical experiments where participants evaluate the realism of the crowd simulation. For example, O'Connor *et al.* [20] observe how social forces of attraction and repulsion determine the perceptual realism of trajectories by varying them in a dynamic context. Barut *et al.* [2] explore how simple collision free rectilinear trajectories generate a plausible crowd motion by creating an illusion of manoeuvring humanoids. In contrast to the aforementioned works, Barut and Haciomeroglu[1] study a simulation with a considerable number of agents (308) simultaneously present in the scene and distributed in an area of approximately 1600 sqm. While these studies explore the perceptual realism of crowd trajectories using 2D images or videos, they do not consider the immersive context. The question therefore of *how being immersed affects the perceived plausibility of crowd trajectories* is not properly addressed in the literature. To the best of our knowledge there does not exist a psychophysical study where participants evaluate crowd trajectories in an immersive context. Our study aims to fill this gap.

2.2 Realism of Crowds in VR

Crowds in VR have been studied extensively in the last decade. Pelechano *et al.* [25] proposed the level of presence achieved by a human as a measure to validate crowds in a VR environment. Olivier *et al.* [23] explore VR as an experimental tool to improve the level of realism of the microscopic pedestrian simulation. In that study VR is used to assess local movements (such as collision avoidance), and trajectories formed by the participants against real situations. Bruneau *et al.* [3] study how participants navigate against the flow of a crowd of simulated characters with different characteristics based on the Principle of Minimum Energy. Another example is [19] which explores human movement behaviour during an immersive virtual crowd interaction. The participant is immersed in an urban context while crossing the road, with different density conditions and directions. They conclude that both direction and speed alter participants' behaviour. A study by Kyriakou *et al.* [15] explores a crowd's perceived realism across different immersion levels (CAVE-like, large 2D display) with crowd interactivity. All these studies focus on the crowd's general plausibility rather than the perception of realism of the crowd's trajectories.

2.3 Point of view in crowd's plausibility/realism

Some of the studies above explore how the point of view (camera position and orientation) impacts a crowd's perceptual realism. Since immersion is a topic highly correlated to the point of observation [30] we consider these studies highly informative. Barut and Haciomeroglu[1] use camera angle as an experimental condition in evaluating the crowd's plausibility; this study, which uses videos, suggests that aerial views increase the perceived effect of pedestrian manoeuvring illusion, thus the perceived crowd's plausibility. Ennis *et al.* [5] explores how the crowds' perceived realism changes across multiple points of static views (eye-level camera vs plan view camera); findings from this study show how, in general, using eye-level views will result in a more realistic representation of static pedestrian crowd formations to a viewer. However, this study only uses static crowds, thus lacks the motion aspect of pedestrian trajectories entirely. Interestingly, both studies agree that the crowd's plausibility changes based on the point of view. Still, they report more realism by exploiting different effects (higher plausibility given by motion illusion or different view points). However Barut and Haciomeroglu[1] consider a scenario in which pedestrians are moving in random directions; they start from random positions in a squared area, rather than a scenario in which pedestrians' flow moves according to the space's features (i.e. walking along a street, moving across the pedestrian crossing, etc..). Thus, we aim to answer the following question: are linear and fixed speed trajectories plausible in a context in which pedestrians need to assess features of the scene?

3 CONCEPT AND MOTIVATION

With this study, we address the challenge of understanding if computationally inexpensive algorithms can generate plausible or even realistic trajectories in a scenario showing pedestrians following urban constraints such as a road crossing. Previous literature [2] suggests that very simple crowd behaviours, such as ones having rectilinear and single speed trajectories can be perceived as plausible while being computationally inexpensive. As discussed in Section2.3, trajectory plausibility is affected by motion illusion and viewpoint. Thus, we aim to clarify if the motion illusion effect can be perceived even in this constrained urban condition where pedestrians are moving in opposite directions. Finally, answering how much the point of view affects the plausibility of such a trajectory generation algorithm is challenging. The observer viewpoint can alter the motion illusion of pedestrians. Understanding the viability of such rectilinear, constant-speed trajectories observed from an immersive point of view will allow us to characterize requirements for crowd trajectory generation algorithms in VR; in particular, the new generation of untethered VR devices based on mobile chipsets, where battery life is a primary concern. Validating a light algorithm for crowd motion without compromising the trajectories' perceived realism can save CPU/GPU time and memory space.

3.1 Realism

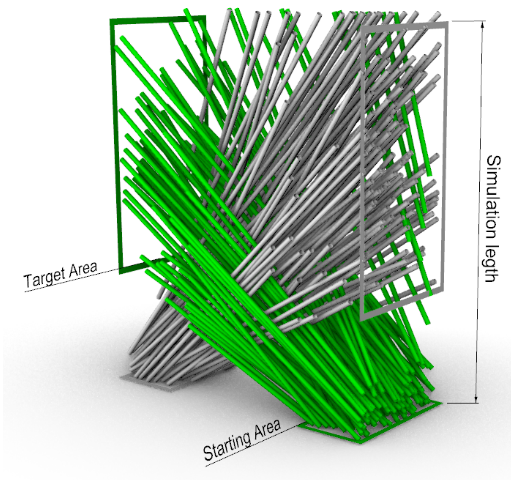
3.1.1 Realism of Trajectories. As the project aims to explore the perceived realism of crowd rectilinear/fix-speed trajectories, we identified a viable baseline. Our baseline consists of real trajectories traced from an aerial video of a pedestrian crossing captured by a drone and containing 280 pedestrians. Trajectories are traced at 25 frames per second by tracing the head of pedestrians. Periodic movements of pedestrians heads are cleaned with a Kalman filter, the resulting output is re-sampled at 72 frames per second which corresponds to the frame rate of the VR headset. The video allows us to trace the pedestrian trajectories and trace the dimension and specifics of the pedestrian crossing scene. The second dataset is subsequently generated using an adaptation of Barut *et al.* [1]. We adapted Barut *et al.* 's real-time rectilinear algorithm to produce trajectories of pedestrians on a crosswalk or pedestrian crossing. We introduced constraints to produce two crowds moving through the pedestrian crossing in opposite directions. This was achieved by creating two starting areas and an equivalent number of target areas (see Fig. 3a). To ensure that the pedestrians' speed distribution was the same as the real data, we used Monte Carlo sampling (Fig. 3b). Both the real-world and simulated data represent the two levels of realism of the stimuli we used for the psychophysical experiment.

Fig. 2. Trajectories were extracted from real data video by tracking pedestrian's heads with an open computer vision software library. At a second stage, trajectories were cleaned from periodic walking movements using a Kalman filter and then re-sampled at 72 frames per second.

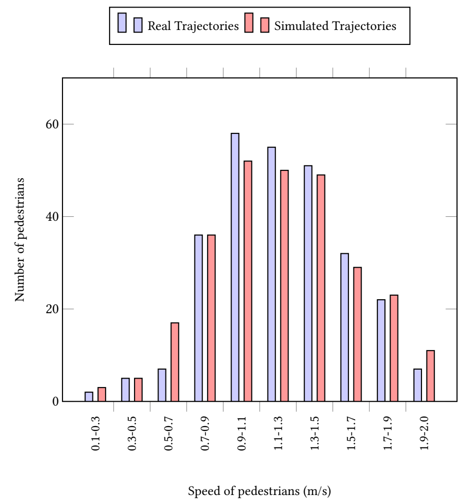
3.1.2 Measuring the perceived realism of trajectories. The previous literature review surrounding perceptual realism of crowds within psychophysical experiments uses a number of different approaches to measure realism. Some employ simple *binary scores* (is the crowd realistic or not) [5] whereas others instead ask participants to asses realism using a *continuous scale* [21] . While the first approach asks participants to make a clear decision, the second allows for the participants to be less deterministic and express their perception with a higher degree of accuracy. Within our questionnaire, we decided to use both approaches. We gather two different levels of perception: a binary classification and a numeric evaluation. Another important aspect of this questionnaire relates to its focus on trajectories. Crowd simulation algorithms focus on generating a trajectory for each member of a crowd. Thus the questionnaire should exclusively collect the participant perception of realism concerning the trajectories rather than the realism of the humanoid movements (i.e. locomotion synthesis) or the realism of the environment. Therefore, the introduction of the questionnaire informs the participant of the focus on trajectories rather than locomotion synthesis.

3.2 Immersion

Immersion is defined as the "illusion" that "the virtual environment technology replaces the user's sensory stimuli with the virtual sensory stimuli" [31]. The realism of the virtual environment is among those characteristics of the experience that influence the sense of immersion [18]. Within this study, we aim to understand to which extent the realism of trajectories impacts the levels of immersion experienced by participants. In order to do so, we use a questionnaire proposed by Tcha *et al.* [31] that aimed to measure the experienced immersion levels. We run this questionnaire immediately after every time a participant experienced one of the experiment conditions. This questionnaire contains 9 questions for which participants are asked to assign a score based on a 10-point Likert scale (1=strongly disagree, 10 = strongly agree).



(a) Rectilinear Trajectories



(b) Speed distribution

Fig. 3. (a) Adaptation of the rectilinear algorithm proposed by [1], in order to simulate a pedestrian crossing. Our adaptation consisted of addding four constraint areas for the generation of the trajectories and the target location. Therefore the algorithm rejects any trajectory which is not compatible with the four constraints. Each of the pipes represents the trajectory of one pedestrian. The pipe radius represents the clearance between pedestrians, and the vertical axis the temporal dimension. (b) The plot shows the average speed distribution of the real data trajectories and the fixed speed distribution of the simulated data trajectories.

4 FIRST EXPERIMENT: PLAUSIBILITY OF RECTILINEAR TRAJECTORIES

The purpose of the first experiment is to validate the plausibility of the rectilinear algorithm proposed by [1] in a realistic scenario. While Barut *et al.* [2] evaluated algorithms in a scenario in which the pedestrian is free to move in any direction, our chosen scenario consists of pedestrians at a road crossing; (i.e., two pedestrian flows in opposite directions). In addition, we aim to clarify if participants perceive the manoeuvring illusion and speed changes that Barut *et al.* report in their study [2]. Our hypothesis is that rectilinear trajectories in such scenario manifest the same perceptual realism experienced in Barut's work. We adopt a within-group experiment design in which all participants are exposed to all experimental conditions consisting of two different points of view: eye and aerial levels (Fig. 4a). During each experimental session, each participant is exposed to two videos from two different points of view and, after seeing each video, they are asked to fill the questionnaire (Table 1). The order of the experimental condition is randomized to remove any learning effect. ANON Ethics Committee approved our experiment. In subsection 4.1 we describe how we generated data for the experiment, then we detail information about the participants in subsection 4.2, the experiment procedure we used in subsection 4.3, the apparatus in subsection 4.4 and finally we discuss the results in subsection 4.5.

	Rectilinear Trajectories Simulation 1	Rectilinear Trajectories Simulation 2
Eye-level View	Condition	Condition
Aerial View	Condition	Condition

(a) Plausibility Experiment

	Real Trajectories	Rectilinear Trajectories
Immersive VR (headset)	Condition	Condition

(b) VR Experiment

(c) Images of experimental conditions

Fig. 4. (a) In the first experiment a total of 4 experimental conditions were defined by the two independent variables of view angle and simulation id; (b) the second experiment only had one independent variable (Real vs Simulated Trajectories) and therefore 2 experimental conditions. (c) From left to right: Aerial view (video experiment 1), Eye level view (video experiment 1), Immersive VR View (3D dynamic scene experiment 2).

4.1 Simulated trajectories

We generated two synthetic simulations that were rendered from two different camera angles. Both synthetic simulations consisted of constant velocity, rectilinear, collision-free paths (see Fig. 3a). In both cases, the entry and exit positions alongside the time that pedestrians appeared in the real-world video (Fig. 2) were used to initialize starting areas, goals area and preferred velocities for each simulated character. For the preferred velocities in particular, a Monte Carlo approach was used; preferred speeds were sampled from the *speed distribution* of the actual pedestrians (Fig. 3b). We also assume that each pedestrian can be represented as a $2D$ circle with a radius $r = .3m$; a straight path can therefore be thought of as a cylinder in 3D space (2D spatial coordinate and time). This representation allows for the specification of a simple geometric based algorithm for crowd simulation. Whenever a new path (cylinder) is added, it is checked for collision with all the existing cylinders; in the case it collides, its speed is resampled, a new cylinder is defined and the process repeats until all the collision free trajectories are generated. Running this algorithm twice gives different results though all simulated results generate rectilinear collision free paths that satisfy the speed distribution of the input data (Fig. 3b).

The generated trajectories were then imported into the Unity Game Engine where Rocketbox characters [6] were used to represent each simulated pedestrian. A custom animation system was used to animate the characters on the trajectories using *different locomotion gaits* without any foot-sliding artifacts. The two simulations were then rendered into videos using two different angles;eye-level and aerial views (Fig. 1).

4.2 Participants

The online test was accessible for 4 weeks and we recruited 153 participants who performed the experiment remotely. The mean age was 32.5 years with standard deviation 10.5 years. The participants comprised of 79 females and 74 males.

4.3 Experiment Procedure

Potential participants were contacted via email, and if interested in taking part, they were provided a link to the questionnaire. Once they arrived on the questionnaire web page, they were given the general study information and the ethics application and asked to consent. Following this they were asked to complete a demographic section. Subsequently a randomized sequence of two videos from two viewpoints were generated. The sequence contained both simulation but viewed from different positions. After viewing each of the videos participants were required to complete the perceived realism questionnaire (Table 1).

Did the pedestrian manoeuvre to avoid collisions? - Yes (I saw pedestrian changing direction to avoid collisions with other pedestrians) - No (I did not see any of the pedestrians changing direction)
Did any of the pedestrians change speed to avoid collisions? - Yes (I saw pedestrian slowing down / speeding up to avoid collisions) - No (I did not see pedestrian changing their speed to avoid collisions)
What is the level of realism of the trajectories? - Realistic (Accurate - I felt pedestrian trajectories resemble real-life trajectories) - Plausible (Credible - I felt pedestrian trajectories were possible/valid but not real) - Implausible (Questionable- I felt pedestrian trajectories were unconvincing) - Impossible (Absurd - I felt pedestrian trajectories were unreasonable)

Table 1. Realism questionnaire

4.4 Experimental Apparatus

We developed an online questionnaire using the PsyToolkit platform [29]. The online questionnaire contained an initial demographics section and then the two randomized conditions of the experiment. All simulations were rendered in 25 frames/sec videos having $1920x1080$ pixels resolution; these were compressed using H.264 compression with constant bitrate and quality settings that minimized compression artifacts and were later uploaded on YouTube.

4.5 Results of Plausibility experiment

Most participants reported the rectilinear algorithm as either realistic or plausible (75% on the eye-level condition and 80% on the aerial view condition). We performed Wilcoxon signed-rank tests to see if the point of view bore any statistical significance regarding the perceived realism of the crowd trajectories. We chose Wilcoxon signed-rank test as our sample distributions were not normal. The Wilcoxon signed-rank test indicated that perceived realism in the aerial view was statistically significantly higher than in the eye-level view ($Z = 2079$, $p < 0.034$, Fig 5a). In addition, the Wilcoxon test indicated that perceived realism for simulation 1 was not statistically significantly different than simulation 2 view ($Z = 717.500$, $p < 0.11$). In the question *"Did the pedestrian manoeuvre to avoid collisions?"* only 28% responded yes to the experimental condition eye-level view, while in the aerial view 50% responded yes. Similar results were reported for the question *"Did any of the pedestrians change speed to avoid collisions"* with 32% answering yes in the eye-level view and 50% in the aerial view. We further analyzed if the point of view implies any statistical significance related to trajectory linearity perception (manoeuvres) and speed changes by performing the Wilcoxon signed-rank test again. Trajectory linearity perception in the aerial view was higher than in the eye-level view ($Z = 1247$, $p < 0.001$, Fig 5b) while Wilcoxon test did not show significant statistical differences in the measures of speed homogeneity across the point of view ($Z = 559$, $p < 0.125$, Fig 5c). Results validate the idea that the rectilinear algorithms generate trajectories that are perceived by most as plausible. Subsequently, it can be used as a plausible synthetic alternative to real data of tracked pedestrians.

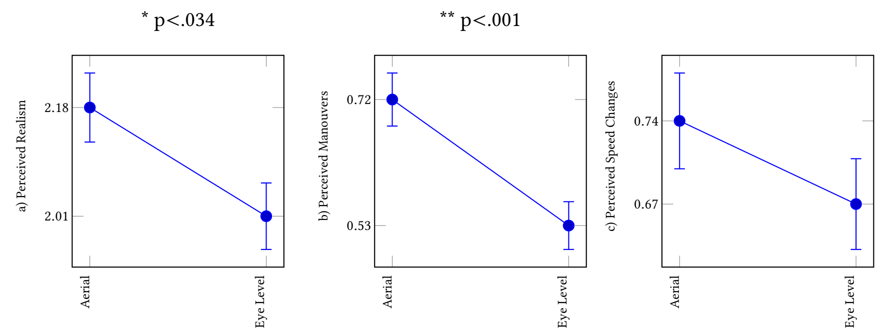


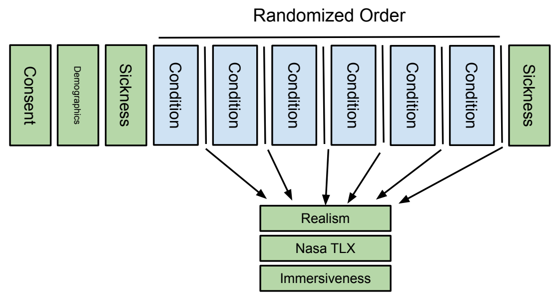
Fig. 5. a) Perceived realism comparison across Aerial View and Eye Level View: on the y axe values of realism between 0 and 3 where 0 is impossible and 3 is realistic. b) Perceived Manoeuvres comparison across Aerial View and Eye Level View. On the y axe value of the perceived maneuvers between 0 and 1 where 0 represent pedestrian moving in straight trajectories and 1 represent pedestrian moving in curved trajectories to avoid collisions. c) Perceived speed changes comparison comparison across Aerial View and Eye Level View. On the y axe value of the perceived maneuvers between 0 and 1 where 0 represent pedestrian moving at a constant speed and 1 represent pedestrian changing speed to avoid collisions. a,b,c) error bars represents Standard Error.

5 SECOND EXPERIMENT: TRAJECTORY PERCEPTION OF CROWDS IN VR

The purpose of this experiment was two-fold: first, we aimed to understand if there was a difference in the perceived realism of a crowd visualised with trajectories acquired from real data compared with synthetic trajectories (linear and constant speed trajectories) when the user is immersed in a virtual environment. Second, we performed a comparison between the participants' estimation of the density and size of the crowd in the same immersive scene. We asked the participants 3 questions related to the realism perception as in the first experiment: firstly, a direct question about the perceived realism (Fig. 9 a). Secondly, if they noticed manoeuvres to avoid collision by the characters (Fig. 9b) and last if the characters speed changed to avoid collision (Fig. 9c). Our experiment was approved by ANON Ethics Committee. We hypothesised that their perception of realism would not differ between real and synthetic trajectories in this immersive condition. In subsection 5.1 we detail information about participants, then we explain experiment procedure in subsection 5.2 and the apparatus in subsection 5.3. The final subsection 5.4 shows and discusses the results.

We adopted a within-group experimental design in which all participants are exposed to both of the realism conditions (i.e. real data and simulated data) as shown in Fig. 4b. This is a task with minimal stress and difficulty levels for the participants since each of the two simulations lasted around 30 seconds. To counterbalance any possible ordering effects, the sequence of conditions is randomized for each participant. Before each experiment, we collected consent, demographics, and a pre-experiment sickness questionnaire. Across every trial, we collected the this realism questionnaire. The real-world trajectories were tracked from a video using the semi-automatic procedure of tracking heads discussed in Section 3.1.1. The synthetic data were generated using the linear trajectories algorithm discussed in Section 4.1.

(a) Participants in controlled environment



(b) Experiment procedure

Fig. 6. (a) Participants from the controlled situation. The device was disinfected after every trial. (b) Experiment procedure.

5.1 Participants

We recruited 40 participants, 15 participants performed the VR experiment remotely, while the remaining 25 took part in the experiment in a controlled environment. For the controlled situation in particular, the device was disinfected after each trial to mitigate against the Covid-19 infection risk. The mean age of participants was 36.2 years and the standard deviation 11.2 years. Among the participants, there were 16 females and 24 males. 7 participants reported not having any experience with VR, 10 participants with low experience, 8 with some experience, 6 with average experience, while 9 reported being very experienced with VR.

5.2 Experiment procedure

Once recruited, remote participants were invited to download the VR application via Side Quest or the ANON Research Network, while in person participants were handed an Oculus Quest with a pre-loaded application. The application was pre-installed on an Oculus Quest 2 for the controlled experiment participants. Once the application was downloaded/installed and opened, a series of graphical instructions (Fig 7) informed the participant about the aim of the project as well as asking for the consent of the participant, which was recorded, before demographic and pre-experiment sickness questionnaires were collected (as shown in Fig. 6b). After the experiment started, randomly selecting which experiment to be shown first, instructions were shown explaining to the participant that they would be asked to observe a crowd and subsequently asked to rate the realism of the crowd trajectories. There was a specific instruction that the participants should focus on the trajectories rather than the humanoid movements or the environment details. After the introductory instructions, the experiment began. The pedestrian crossing crowd scene lasted 28 seconds, and after it ended, the participant could choose to observe it for a second time. Between each trial, participants were asked a number of questions related to perceived realism. At the end of the experiment, participants were asked to fill a post-experiment sickness questionnaire [13]. The experiment typically lasted less than 10 minutes.

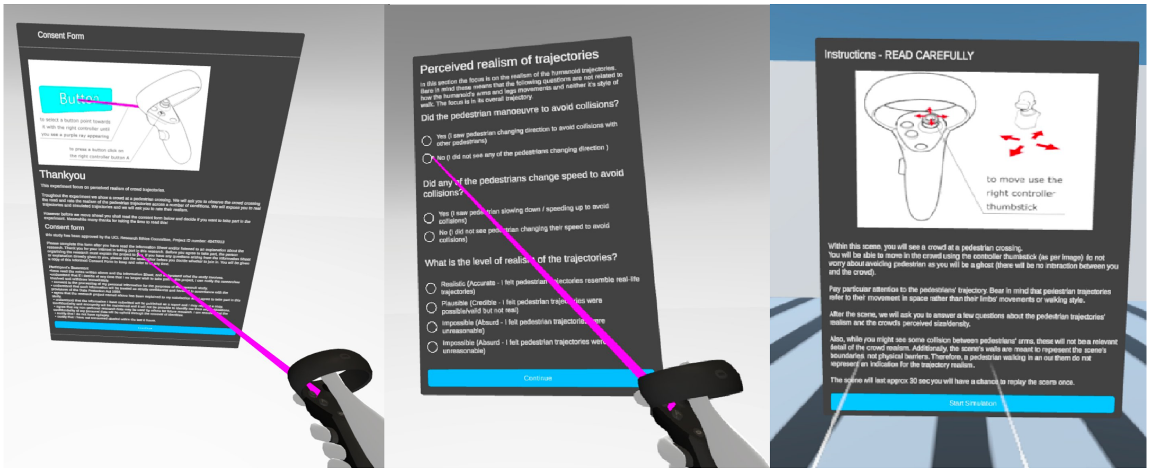


Fig. 7. Graphical instruction and form of the VR experiment.

5.3 Experimental Apparatus

Our experiment was designed to run unsupervised. After having run the application, the user followed the instructions displayed in VR. The VRCrowd application was developed with Unity 2019.4.0 and uses Google Firebase to collect the data. A network connection therefore needed to be present to run the experiment. Participants needed to stand in an area that allowed them to orientate towards different directions. No locomotion in the real environment was required as user can navigate in virtual environment via controller.

5.3.1 Trajectories. We showed two sets of trajectories: one set consisting of the real trajectories traced from the aerial video (refer to Section 3.1.1), and another set of generated trajectories (as explained in Section 4.1).

5.3.2 Locomotion synthesis. To increase the realism of the model animation and model visual appearance, we animated Rocketbox characters [6] using the phase functional neural networks animation system developed by Holden *et al.* [9]. For both the real and the artificial trajectories, we firstly applied the animations offline by calculating the trajectories with Unity. Then we recorded and stored the dataset in binary files to be played within the Oculus. This solution allowed us to run the state-of-the-art walking animation algorithm on more than 250 humanoids into a VR standalone device, with a fluid animation even in a standalone VR device. To avoid possible slowdowns caused by the rendering loop, we implemented a LOD system for the humanoids. To match Rocketbox rigged models with the exact skeleton orientations and positions from Holden's research and transfer the body movements correctly, we used the Unity Animation Rigging package.

5.4 Results

We did not measure significant differences between real and synthetic data in the results from both the questions related to collision avoidance and speed changes

While participants were undecided if the pedestrians changed trajectories to avoid collisions in the first case, in the second 70% of the participants reported no speed changes. To see if there were any statistical differences in the sample distribution of these answers, we performed the Wilcoxon signed-rank test. The results of this test showed that the

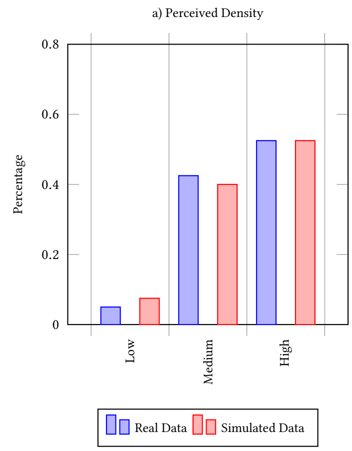


Fig. 8. This chart shows the comparison between the perceived crowd density between real and simulated data in the second experiment. In both cases the participants perceive the crowd density between medium and large.

perception of manoeuvring in the real data condition was not significantly different (Z=0, p<.346, Fig. 9b) across the real/simulated data. This was also the case for perceived speed changes (Z=0,p=1, Fig. 9c). The lack of a significant difference between perceptions of simulated and real data suggests that *when the user is immersed, rectilinear trajectories are perceived as similarly realistic as real trajectories.* In the real world data, pedestrians change both speed and direction to avoid collisions whereas in the simulated data, pedestrians walk on precomputed, collision-free paths with constant speeds, without changing their direction of movement.

Similar results were also visible in the distribution of answers for the perceived realism questions.To see if participants perceived the real trajectories differently compared with the simulated trajectories, we performed Wilcoxon signed rank test. The Wilcoxon signed-rank test indicated that *the perceived realism of real data was not statistically significantly different than in the eye-level view* (Z= 6, p<.149, Fig. 9a).

Participants were also asked to estimate two specific features of the crowds; *density* and *size* (number of pedestrians). We provided three possible values for the crowd density; low, medium or high. From Nelson [19] we defined low crowd density as 1 agent per square meter, medium density as 1.5 agents per square meter and high density as 2 agents per squared meter (Fig 8).

In the case of crowd size we provided several ranges as options. We compared the perceived crowd density and perceived crowd size between real data and simulated data by performing again the Wilcoxon signed rank test. We noticed that participants indicated high or medium, with a cumulative percentage of more than 90% for both simulated and real data. Perception of crowd density in the real data condition was not significantly different from the simulated data (Z=1, p=1 Fig. 9d). The results were similar for the perceived crowd size, which show no statistically significant differences across the real/simulated data Z=11.500 p=0.322 (Fig. 9e).

Finally, we aim to understand if the trajectory realism affects the level of immersion experienced by participants. We do so by comparing the scores distributions of the immersion questionnaire [31] to see if there is any statistically significance difference across the two samples. The Wilcoxon signed-rank test indicates that immersion level reported by participants across the two levels is not statistically significantly different (Z= 376, p<.656, Fig. 9f).

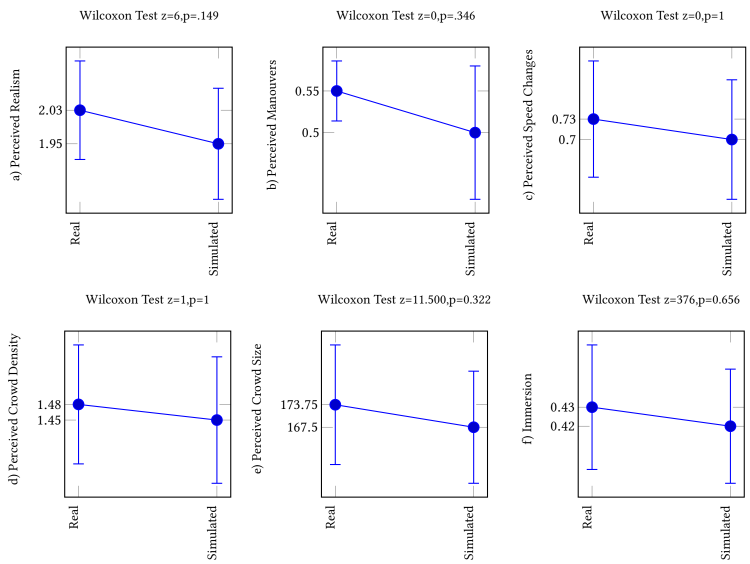


Fig. 9. a) Perceived realism comparison across Real and Simulated trajectories: on the y axes values of realism between 0 and 3 where 0 is impossible and 3 is realistic. b) Perceived Manoeuvres comparison across Real and Simulated trajectories: on the y axes value of the perceived maneuvers between 0 and 1. Where 0 represent perception of pedestrian moving in straight trajectories and 1 represent perception of pedestrian moving in curved trajectories to avoid collision. c) Perceived speed changes across Real and Simulated trajectories: on the y axes value of the perceived speed changes between 0 and 1. 0 represent perception of pedestrian moving at a fix speed and 1 represent perception of pedestrian changing speed to avoid collision. d) Perceived crowd density across Real and Simulated trajectories. On the y axes value of the perceived crowd density between 0 and 2. 0 represent low density, 1 represent medium density and 2 represent high density. e) Perceived crowd size across Real and Simulated trajectories. On the y axes the estimated number of pedestrians. f) Reported level of immersion across real and simulated trajectories. On the y axes the score of the immersion questionnaire proposed and validated by Tcha *et al.* [32]. The error bars in all the sub-charts represent Standard Error.

6 LIMITATIONS AND FUTURE WORK

The first experiment (Section 4) evaluates the perceived realism of rectilinear trajectory crowd simulations on 2D displays. Such trajectories are subsequently used in the second experiment that focuses on the perception of the realism of the trajectories performed by the characters in a crowd simulation in VR (Section 5). Our experiment positions the observer in the same environment as two pedestrian flows crossing a road. Despite the scenario's specificity, this study aims to investigate a common configuration when crowd simulations are evaluated. We aim to extend the study by increasing the number of different situations such as singular flow, four flows in a crossroad context, and bottleneck scenarios. A further improvement is to increase the number of persons in the crowd to have a high-density simulation in a larger environment, still matching the real-time requirements on a VR device.

7 CONCLUSION

Results from the first experiment highlight how the adapted algorithms from [1] can be used to generate plausible/realistic trajectories even in conditions in which the flow of pedestrians is not random such as in the case of a pedestrian crossing. The results also confirm how the aerial viewpoint leads to a more realistic perception of trajectories than the eye-level viewpoint, confirming that the point of view affects the realism of rectilinear trajectories.

The results from the second experiment highlight how, in an immersive condition, we can not determine differences in perceived realism between traced trajectories and simulated trajectories. This promising outcome suggests that spending computational power for trajectory realism does not necessarily increase the perceived realism in a crosswalk setting with a high density crowd. To support this, the results confirm that participants perceive pedestrian manoeuvring to avoid collisions even if pedestrians move in rectilinear directions, even when they observe them from an immersive viewpoint. Additionally, they perceived that pedestrians were changing their speed to avoid collisions. In all cases, the difference in perception of pedestrian manoeuvres, the speed changes and ultimately the overall realism of trajectories is not statistically different across the two levels of realism. These suggest that participants cannot consistently distinguish between realistic trajectories (actually not always rectilinear and subjected to speed variations) and rectilinear, constant speed trajectories within the immersive condition. These experiments validate the rectilinear trajectory generation as a viable approach to animate high-density crowds in VR.

REFERENCES

[1] Oner Barut and Murat Haciomeroglu. 2015. Real-time collision-free linear trajectory generation on GPU for crowd simulations. *Visual Computer* 31, 6-8 (2015), 843–852. https://doi.org/10.1007/s00371-015-1105-z

[2] Oner Barut, Murat Haciomeroglu, and Ebru Akcapinarsezer. 2018. Perceptual evaluation of maneuvering motion illusion for virtual pedestrians. *Visual Computer* 34, 6-8 (6 2018), 1119–1128. https://doi.org/10.1007/s00371-018-1557-z

[3] Julien Bruneau, Anne-Hélène Olivier, and Julien Pettré. 2015. Going Through, Going Around: A Study on Individual Avoidance of Groups. *IEEE transactions on visualization and computer graphics* 21, 4 (April 2015), 520–528. https://doi.org/10.1109/TVCG.2015.2391862

[4] Panayiotis Charalambous, Ioannis Karamouzas, Stephen J. Guy, and Yiorgos Chrysanthou. 2014. A Data-Driven Framework for Visual Crowd Analysis. *Computer Graphics Forum* 33, 7 (Oct. 2014), 41–50. https://doi.org/10.1111/cgf.12472

[5] Cathy Ennis, Christopher Peters, and Carol O'sullivan. 2011. Perceptual effects of scene context and viewpoint for virtual pedestrian crowds. *ACM Transactions on Applied Perception* 8, 2 (1 2011), 1–22. https://doi.org/10.1145/1870076.1870078

[6] Mar Gonzalez-Franco, Eyal Ofek, Ye Pan, Angus Antley, Anthony Steed, Bernhard Spanlang, Antonella Maselli, Domna Banakou, Nuria Pelechano, Sergio Orts-Escolano, Veronica Orvalho, Laura Trutoiu, Markus Wojcik, Maria V. Sanchez-Vives, Jeremy Bailenson, Mel Slater, and Jaron Lanier. 2020. The Rocketbox Library and the Utility of Freely Available Rigged Avatars. *Frontiers in Virtual Reality* 1, November (2020), 1–23. https://doi.org/10.3389/frvir.2020.561558

[7] Stephen J. Guy, Jatin Chhugani, Sean Curtis, Pradeep Dubey, Ming C. Lin, and Dinesh Manocha. 2010. PLEdestrians: A Least-Effort Approach to Crowd Simulation. In *Proceedings of the 2010 Eurographics/ACM SIGGRAPH Symposium on Computer Animation, SCA 2010, Madrid, Spain, 2010*, Zoran Popovic and Miguel A. Otaduy (Eds.). Eurographics Association, 119–128. https://doi.org/10.2312/SCA/SCA10/119-128

[8] Stephen J. Guy, Jur Van Den Berg, Wenxi Liu, Rynson Lau, Ming C. Lin, and Dinesh Manocha. 2012. A statistical similarity measure for aggregate crowd dynamics. In *ACM Transactions on Graphics*, Vol. 31. 1–11. https://doi.org/10.1145/2366145.2366209

[9] Daniel Holden, Taku Komura, and Jun Saito. 2017. Phase-functioned neural networks for character control. *ACM Transactions on Graphics* 36, 4 (2017), 1–13. https://doi.org/10.1145/3072959.3073663

[10] Ludovic Hoyet, Anne Hélène Olivier, Richard Kulpa, and Julien Pettré. 2016. Perceptual effect of shoulder motions on crowd animations. *ACM Transactions on Graphics* 35, 4 (2016). https://doi.org/10.1145/2897824.2925931

[11] Mubbasir Kapadia, Shawn Singh, Brian Allen, Glenn Reinman, and Petros Faloutsos. 2009. SteerBug: an interactive framework for specifying and detecting steering behaviors. In *Proceedings of the 2009 ACM SIGGRAPH/Eurographics Symposium on Computer Animation - SCA '09*. ACM Press, New Orleans, Louisiana, 209. https://doi.org/10.1145/1599470.1599497

[12] Ioannis Karamouzas, Brian Skinner, and Stephen J. Guy. 2014. Universal Power Law Governing Pedestrian Interactions. *Phys. Rev. Lett.* 113 (Dec 2014), 238701. Issue 23. https://doi.org/10.1103/PhysRevLett.113.238701

[13] Robert S. Kennedy, Norman E. Lane, Kevin S. Berbaum, and Michael G. Lilienthal. 1993. Simulator Sickness Questionnaire: An Enhanced Method for Quantifying Simulator Sickness. *The International Journal of Aviation Psychology* 3, 3 (1993), 203–220. https://doi.org/10.1207/s15327108ijap0303{_}3

[14] Meeko Kuwahara, Takao Sato, and Yuko Yotsumoto. 2012. Wriggling motion trajectory illusion. *Journal of Vision* 12, 12 (2012), 1–14. https://doi.org/10.1167/12.12.4

[15] Marios Kyriakou, Xueni Pan, and Yiorgos Chrysanthou. 2017. Interaction with virtual crowd in Immersive and semi-Immersive Virtual Reality systems. *Computer Animation and Virtual Worlds* 28, 5 (2017), 1–12. https://doi.org/10.1002/cav.1729

[16] Alon Lerner, Yiorgos Chrysanthou, Ariel Shamir, and Daniel Cohen-Or. 2010. Context-Dependent Crowd Evaluation. *Computer Graphics Forum* 29, 7 (Sept. 2010), 2197–2206. https://doi.org/10.1111/j.1467-8659.2010.01808.x

[17] Artur Malinowski, Paweł Czarnul, Krzysztof Czuryło, Maciej Maciejewski, and Paweł Skowron. 2017. Multi-agent large-scale parallel crowd simulation. *Procedia Computer Science* 108 (Jan. 2017), 917–926. https://doi.org/10.1016/j.procs.2017.05.036

[18] Rory McGloin, Kirstie M. Farrar, and Marina Krcmar. 2011. The impact of controller naturalness on spatial presence, gamer enjoyment, and perceived realism in a tennis simulation video game. *Presence: Teleoperators and Virtual Environments* 20, 4 (2011), 309–324. https://doi.org/10.1162/PRES{_}a{_}00053

[19] Michael G. Nelson, Alexandros Koilias, Sahana Gubbi, and Christos Mousas. 2019. Within a virtual crowd: Exploring human movement behavior during immersive virtual crowd interaction. In *Proceedings - VRCAI 2019: 17th ACM SIGGRAPH International Conference on Virtual-Reality Continuum and its Applications in Industry*. Association for Computing Machinery, Inc, New York, NY, USA, 1–10. https://doi.org/10.1145/3359997.3365709

[20] Stuart O'Connor, Fotis Liarokapis, and Chrisina Jayne. 2015. Perceived realism of crowd behaviour with social forces. In *Proceedings of the International Conference on Information Visualisation*, Vol. 2015-Septe. Institute of Electrical and Electronics Engineers Inc., 494–499. https://doi.org/10.1109/iV.2015.88

[21] Stuart O'Connor, Fotis Liarokapis, and Christopher Peters. 2013. An initial study to assess the perceived realism of agent crowd behaviour in a virtual city. In *2013 5th International Conference on Games and Virtual Worlds for Serious Applications, VS-GAMES 2013*. https://doi.org/10.1109/VS-GAMES.2013.6624220

[22] Stuart O'Connor, James Shuttleworth, Simon Colreavy-Donnelly, and Fotis Liarokapis. 2019. Assessing the perceived realism of agent grouping dynamics for adaptation and simulation. *Entertainment Computing* 32, September (2019), 100323. https://doi.org/10.1016/j.entcom.2019.100323

[23] Anne Hélène Olivier, Julien Bruneau, Gabriel Cirio, and Julien Pettré. 2014. A virtual reality platform to study crowd behaviors. *Transportation Research Procedia* 2 (2014), 114–122. https://doi.org/10.1016/j.trpro.2014.09.015

[24] Nuria Pelechano, Jan M. Allbeck, Mubbasir Kapadia, and Norman I. Badler. 2016. *Simulating Heterogeneous Crowds with Interactive Behaviors*. CRC Press, Taylor & Francis Group, 6000 Broken Sound Parkway NW, Suite 300, Boca Raton, FL 33487-2742. https://doi.org/10.1201/9781315370071

[25] Nuria Pelechano, Catherine Stocker, Jan Allbeck, and Norman Badler. 2008. Being a part of the crowd: Towards validating VR crowds using presence. *Proceedings of the International Joint Conference on Autonomous Agents and Multiagent Systems, AAMAS* 1, May (2008), 135–141. https://doi.org/10.1145/1402383.1402407

[26] P. S. A. Reitsma and N. S. Pollard. 2004. Evaluating motion graphs for character navigation. In *Proceedings of the 2004 ACM SIGGRAPH/Eurographics symposium on Computer animation - SCA '04*. ACM Press, Grenoble, France, 89. https://doi.org/10.1145/1028523.1028536 ISSN: 17275288.

[27] Alastair Shipman and Arnab Majumdar. 2018. Fear in Humans: A Glimpse into the Crowd-Modeling Perspective. *Transportation Research Record: Journal of the Transportation Research Board* 2672, 1 (12 2018), 183–197. https://doi.org/10.1177/0361198118787343

[28] Shawn Singh, Mubbasir Kapadia, Petros Faloutsos, and Glenn Reinman. 2009. SteerBench: a benchmark suite for evaluating steering behaviors. *Computer Animation and Virtual Worlds* 20, 5-6 (Sept. 2009), 533–548. https://doi.org/10.1002/cav.277

[29] Gijsbert Stoet. 2010. PsyToolkit: A software package for programming psychological experiments using Linux. *Behavior research methods* 42, 4 (2010), 1096–1104.

[30] Laurie N Taylor. 2002. Video games: Perspective, point-of-view, and immersion. *Vasa* April (2002), 35. http://purl.fcla.edu/fcla/etd/UFE1000166

[31] Katy Tcha-Tokey, Olivier Christmann, Emilie Loup-Escande, and Simon Richir. 2016. Proposition and Validation of a Questionnaire to Measure the User Experience in Immersive Virtual Environments. *International Journal of Virtual Reality* 16, 1 (2016), 33–48. https://doi.org/10.20870/ijvr.2016.16.1.2880

[32] Katy Tcha-Tokey, Emilie Loup-Escande, Olivier Christmann, and Simon Richir. 2016. A questionnaire to measure the user eXperience in immersive virtual environments. In *ACM International Conference Proceeding Series*. Association for Computing Machinery, New York, New York, USA, 1–5. https://doi.org/10.1145/2927929.2927955

[33] Wouter van Toll, Atlas F. Cook IV, and Roland Geraerts. 2012. Real-time density-based crowd simulation. *Comput. Animat. Virtual Worlds* 23, 1 (2012), 59–69. https://doi.org/10.1002/cav.1424

[34] Guillermo Vigueras, Juan M. Orduña, Miguel Lozano, and Yiorgos Chrysanthou. 2011. A distributed visualization system for crowd simulations. *Integrated Computer-Aided Engineering* 18, 4 (Jan. 2011), 349–363. https://doi.org/10.3233/ICA-2011-0381 Publisher: IOS Press.

[35] He Wang, Jan Ondřejy, and Carol O'Sullivanz. 2016. Path Patterns: Analyzing and comparing real and simulated crowds. *Proceedings - 20th ACM SIGGRAPH Symposium on Interactive 3D Graphics and Games, I3D 2016* (2016), 49–58. https://doi.org/10.1145/2856400.2856410

[36] D. Wolinski, S. J. Guy, A. H. Olivier, M. Lin, D. Manocha, and J. Pettré. 2014. Parameter estimation and comparative evaluation of crowd simulations. *Computer Graphics Forum* 33, 2 (2014), 303–312. https://doi.org/10.1111/cgf.12328