



# Eye-Gaze Activity in Crowds: Impact of Virtual Reality and Density

Florian Berton, Ludovic Hoyet, Anne-Hélène Olivier, Julien Bruneau, Olivier Le Meur, Julien Pettré

## ► To cite this version:

Florian Berton, Ludovic Hoyet, Anne-Hélène Olivier, Julien Bruneau, Olivier Le Meur, et al.. Eye-Gaze Activity in Crowds: Impact of Virtual Reality and Density. VR 2020 - 27th IEEE Conference on Virtual Reality and 3D User Interfaces, Mar 2020, Atlanta, United States. pp.1-10. hal-02544516

**HAL Id: hal-02544516**

**<https://hal.science/hal-02544516>**

Submitted on 16 Apr 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Eye-Gaze Activity in Crowds: Impact of Virtual Reality and Density

Florian Berton\*   Ludovic Hoyet\*   Anne-Hélène Olivier\*   Julien Bruneau\*   Olivier Le Meur\*  
Julien Pettre\*

Univ Rennes, Inria, CNRS, Irisa, M2S, France



Figure 1: Our objective is to analyze eye-gaze activity within a crowd to better understand walkers' interaction neighborhood and simulate crowd behaviour. We designed 2 experiments where participants physically walked both in a real and virtual street populated with other walkers, while we measured their eye-gaze activity (red circle). We evaluated the effect of virtual reality on eye-gaze activity by comparing real and virtual conditions (left and middle-right) and investigated the effect of crowd density (right).

## ABSTRACT

When we are walking in crowds, we mainly use visual information to avoid collisions with other pedestrians. Thus, gaze activity should be considered to better understand interactions between people in a crowd. In this work, we use Virtual Reality (VR) to facilitate motion and gaze tracking, as well as to accurately control experimental conditions, in order to study the effect of crowd density on eye-gaze behavior. Our motivation is to better understand how interaction neighborhood (i.e., the subset of people actually influencing one's locomotion trajectory) changes with density. To this end, we designed two experiments. The first one evaluates the biases introduced by the use of VR on the visual activity when walking among people, by comparing eye-gaze activity while walking in a real and virtual street. We then designed a second experiment where participants walked in a virtual street with different levels of pedestrian density. We demonstrate that gaze fixations are performed at the same frequency despite increases in pedestrian density, while the eyes scan a narrower portion of the street. These results suggest that in such situations walkers focus more on people in front and closer to them. These results provide valuable insights regarding eye-gaze activity during interactions between people in a crowd, and suggest new recommendations in designing more realistic crowd simulations.

**Keywords:** Gaze Activity, Locomotion, Crowd, Virtual Reality, Eye-tracking, Collision Avoidance

**Index Terms:** Human-centered computing—Visualization—Visualization techniques—Treemaps; Human-centered computing—Visualization—Visualization design and evaluation methods

## 1 INTRODUCTION

In this work, we leverage the power of Virtual Reality (VR) to study human locomotion in dynamic environments, for the purpose of modeling and simulating virtual crowds. Modeling crowds requires to design numerical models of local interactions which define by whom and how each individual (or agent) is influenced by the movement of others. The question of the who is also known as the *interaction neighborhood* [57]. Several models have been proposed for this neighborhood, such as a fixed number of nearest agents [59],

or any agents closer than a distance threshold [21], but these solutions were arbitrarily designed based on the study of trajectories. However, it was recently demonstrated that a combined analysis of gaze and trajectory data is meaningful for exploring these questions, e.g., that agents with a high risk of collision are more gaze at [36]. Thus, it is important to develop experimental protocols seeking to simultaneously study gaze and motion data. As such experiments remain difficult to perform in real conditions, VR therefore presents unique opportunities for such studies.

In this paper, our objective is to further understand how eye-gaze activity is influenced by the number of people we are interacting with in a crowd, as a mean of better understanding interaction neighborhood as well as collision avoidance manoeuvres. More precisely, as gaze is indicative of how people take into account other individuals to navigate in crowds [36], we wonder how eye-gaze activity features will be influenced by the density of the crowd. Does the frequency with which people observe other individuals increase with density, enabling them to take a larger number of neighbours into account to adjust their trajectory? Or, on the contrary, is the number of neighbours taken into account constant over time, while people pay more attention to those presenting the greatest risk of collision? To answer these questions, we conducted two VR experiments precisely controlling the visual neighbourhood of an immersed walker, while simultaneously measuring his/her movement and eye-gaze activity:

1. In a first experiment, we estimate the biases induced by the use of VR on eye-gaze activity. We asked participants to walk an existing street, and recorded both their eye-gaze activity and their visual environment. We then reproduced the same situation in VR using a digital replica of the street. We compared the eye-gaze activity of participants under these two conditions. Results show a strong similarity in the eye-gaze activity, while highlighting some quantitative changes that we attribute to the difference in the eye-tracking devices we used as well to the differences in the visual content in the real and virtual conditions.
2. In a second experiment, we evaluate the influence of crowd density on eye-gaze activity. We asked participants to navigate a virtual street populated with different densities of virtual characters. Results show an influence of density on gaze deployment, where participants look more at the center of their visual field as density increases, tending to observe more passers-by in front of them while scanning frequency remains identical.

Our contribution is therefore twofold. We are contributing to the

\*e-mail: firstname.lastname@inria.fr

validation of VR as a tool for studies coupling eye-gaze activity and navigation, showing important similarities of virtual compared to real behaviours. We also propose new ways to improve crowd simulation algorithms by improving knowledge about how the interaction neighborhood of walkers might be visually evaluated by viewers.

## 2 RELATED WORK

### 2.1 Kinematics of interactions between walkers

#### 2.1.1 Pairwise interactions

There is a growing interest in the literature about interactions between walkers, especially focusing on how two walkers avoid each other. It was first studied through the lens of kinematic analysis trajectories performed in real conditions, showing that collision avoidance adjustments are performed only when walkers are on a collision course [41]. More precisely, they showed that walkers trigger motion adaptations only if the future distance of closest approach is below 1m [41]. These adjustments are done by changing speed or orientation of the walking trajectories [3, 23, 40] and are influenced more by situational factors such as crossing angle or crossing order, rather than personal factors such as personality or gender [28].

Those studies performed in real conditions faced the difficulty of reproducing the same stimulus for each participants. Therefore, several studies [2, 9, 39, 41] have been conducted to evaluate the differences in the manoeuvres to perform a collision avoidance with a static or dynamic virtual human in a real and virtual environment. These experimental studies used various VR setups such as an HMD and a CAVE, as well as a multitude of locomotion techniques (joystick, physically walk,...). These studies converged to the same conclusion: virtual reality is a relevant tool to study the kinematics of collision avoidance between walkers. It preserves the nature of motion adaptations but some quantitative differences should be considered. They may be explained by misperception of distances observed in virtual reality [31, 45]. Virtual reality experiments were then recently designed to investigate the information extracted from the motion of the walker to avoid (global vs. local motion cues) [33] as well as the effect of eye contact on collision avoidance behaviour [32, 38].

#### 2.1.2 Multiple interactions

In the context of crowd simulation, it is important to understand more than pairwise interactions only, but as well how these interactions combine and what is the interaction neighbourhood of the walker. In other terms, who influences one's motion when navigating in a crowded situation. In a real environment, Dicks *et al.* [17] designed an experiment where participants had to avoid one or two oncoming walkers. They showed that participants took longer to complete the task when they avoid a collision with two walkers. Meerhof *et al.* [37] proposed another experimental approach, comparing dyadic (1vs.1) and tryadic (1vs.2) situations of collision avoidance in a 90° crossing setup. Results showed that tryadic situations can result both in sequential or simultaneous interactions, and that additional work is needed to identify the conditions which invite for such interactions when multiple walkers are involved. Rio and *al.* [48] studied the behaviour of a participant within a group of virtual walkers whose heading and speed were manipulated. In such a situation, the influence of neighbour was described as a linear function of distance and does not depend on the eccentricity of the other walkers within the participants' field of view.

In close relation to our present research topic, several studies also investigate the effect of density on walker behaviours. One concept developed was named "fundamental diagram" [51] and characterized the relation between speed (or flow) and density in self-organized pedestrian motions. Authors showed a decrease of walking speed with the increase of density, which is also influenced by cultural factors [12]. Bruneau *et al.* [8] showed that the decision of going through or around a group of virtual walkers is influenced

by group density. Using a critical threshold of density to guide the decision to avoid a group of walkers, they proposed an adaptation of RVO model [62] to take into account the presence of groups in a crowd simulation. Finally, in a VR experiment, Dickinson *et al.* recently reported that high crowd density has a negative influence on the affective state of participants, where the task was perceived as uncomfortable [16]. Authors also reported more direction changes and stops in the case of high density levels.

These previous studies provided us with interesting findings to understand how walkers interact with each other but several questions regarding the definition of interaction neighborhood remain unclear. In particular, while it is possible to analyse the trajectory performed by the walker varying the conditions of interaction, it is challenging to define who in the crowd was responsible for the motion adaptations observed. To go further in the analysis, we believe that the study of gaze behaviour would provide relevant insight. Indeed, vision is fundamental in the control of locomotion and it was shown that gaze is directed towards the elements of the environment which maximize the level of information to navigate safely [34]. In a steering task, previous works demonstrated that gaze anticipates the change of direction of walking to collect information about the future direction of motion [5], and this is also true in VR [7]. The following section will present first the definitions and methods in relation to the measure of gaze activity and then the studies investigating gaze behaviour of a walker interacting with their environment.

### 2.2 Gaze activity and interactions while walking

Eye trackers are devices recording the positions of eyes over time, which is used to characterize the gaze behaviour. The gaze behaviour can be described as a succession of fixations which last for about 200 – 300ms, separated by fast eye movement called saccades (30 – 50ms) [42]. Depending on the field of application, different measurements (e.g., duration, amplitude, spatial distribution) can be taken from these variables to study eye activity [29].

Gaze tracking data is used to understand how human interact with their environment, as visual attention reveals some mechanisms to process visual information [35]. For instance, a specific task requires specific information and lead to specific gaze activity patterns [61]. When walking, gaze is attracted by zones which maximize the level of information that can be used to navigate safely [34]. Cinelli *et al.* [13] observed participants going through 2 motor-driven sliding doors, and concluded that gaze fixations depend on the complexity of door movements. Few studies considered collision avoidance between walkers. Kitazawa and Fujiyama [27] studied the relationship between gaze and the Personal Space and observed that gaze allocation was equally distributed between ground, objects and pedestrians. Croft *et al.* [15] studied avoidance strategies between two participants with different velocities, paths and gaze behaviour conditions and found that they predict crossing order. Finally, Jovancevic-Misic and Hayhoe [25] demonstrated that gaze strategies depend on the behaviour of surrounding people, where participants typically looked more at near actors displaying risky behaviours than at other actors.

The integration of eye-tracking capabilities in VR devices such as HMDs greatly facilitates studies on gaze activity. For instance, several studies [11, 56] analysed how visual cues displayed by multiple agents in a crowd affect the gaze of another walker. In particular, they demonstrated that a shared gaze from at least two persons could lead to joint attention with another walker encountered. In a different context, Jovancevic *et al.* [24] asked participants to walk in VR among a few virtual humans (VH) and studied the distribution of gaze fixations in the environment depending on the nature of interactions with VHs, i.e. they focus on following rather than on avoiding. More recently, Meerhoff *et al.* [36] demonstrated that gaze is attracted toward pedestrians with the highest risk of collision when walking in a virtual crowd. However, as the number of such VR studies increases, it also becomes necessary to evaluate the biases

possibly induced on gaze activity by the use of VR. Similar gaze behaviors were found during experiments conducted in both virtual and real environments where participants sat on a chair and observed either a realistic avatar [49] or a light [44], despite differences in head rotations [44]. Same conclusions were reached in a recent study [6], where participants had to avoid another pedestrian while walking in either a virtual or a real environment.

In conclusion, despite its relevancy to provide additional knowledge on interaction neighbourhood, very few studies were conducted on the analysis of gaze activity in virtual crowds. In the present paper, we are interested in gaze movements performed by a participant walking through a crowd of virtual humans. We are more specifically interested in the effect of the level of density on the gaze, i.e., how the crowd density will impact the spatial and temporal distributions of the fixations and the gaze pattern. Furthermore, there is still a lack of work dealing with the bias induced by VR on gaze activity, especially for complex and dynamic situations. This observation allows us to establish our objectives as detailed below.

### 3 OVERVIEW

Our objective is to explore and further understand the interaction neighborhood of people walking in busy environments, with the particular interest of relying on the analysis of the walker's eye-gaze activity. We choose to perform this study in VR, to facilitate the control of experimental conditions, the replication over several participants, as well as the measure of the eye-gaze activity. To this end, we conducted two experiments, the first one allowed us to study the bias induce by VR on eye-gaze activity (Section 4). The second experiment focused on the impact of crowd density on eye-gaze activity (Section 5). We decided to carry out these experiments based on the task of walking in a busy street. The advantage of using such a task is to correspond to a daily-life situation, with no ambiguity on how to realize it: participants simply have to walk and to follow the direction of the street as they commonly do. Having a clear and simple task is important to us, as we know that the nature of the task has a direct impact on the eye-gaze activity [61].

#### 3.1 Apparatus & Task

Participants walked the real, or digital reproduction, of Vasselot street, in the city of Rennes, France (see Figure 1). The digital reproduction was designed by Archivideo, with professional centimetric geometrical precision and textures generated from real photos. Slight differences between the RE and VE were however still present, due to minor differences in the exact localization or aspect of some objects, such as chairs at the terraces of cafés, billboards, etc. In both RE and VE, we were interested in recording participants' eye-gaze activity while they interacted with other pedestrians in the street:

- Real Environment (RE): participants wore in the Tobii pro glasses 2 eye-tracking, which recorded both their eye-gaze activity (50Hz, 4 eye cameras) and a video of their visual field (scene camera: 25Hz, 90° field of view, H.264 1920x1080 pixels)
- Virtual Environment (VE): participants were immersed in the VE using a Fove HMD (70Hz, 100° FoV), which comes with an integrated eye-tracker (100Hz). The virtual scene was rendered using Unity. Participants freely moved in a physical space (gymnasium) of 20m × 6m, while their position was tracked with a 23-camera motion capture system (Qualisys).

In both RE or VE, participants were asked to navigate in a street while avoiding collision with pedestrians and to stop when they were in front of a specific shop. They had to perform multiple round trips between two specific shops (separated by 20m). Figure 1-middle-left) gives an example of the virtual conditions, where participants were asked to navigate in the virtual street by walking in the gymnasium.

For the virtual condition, the virtual humans were driven by RVO [55], an open-source crowd simulator often used in video-games [54]. Its computational performances enable to have multiple agents avoiding collisions with other obstacles without impacting the framerate, which is crucial for VR experiments. In our experiments, RVO parameters were the following: each agent was represented by a 0.5m-radius cylinder, took into account a maximum of 7 neighbours in a 5m space around them, was assigned a random speed  $\in [0.95, 1.25]m/s$ , and were set up to perform collision-avoidance manoeuvres 3s before a potential collision. We chose a distribution centered around 1.1m/s instead of 1.3m/s for the agent's comfort speed as participant are walking slower in VR [1, 10].

#### 3.2 Participants

Twenty-one unpaid participants, recruited via internal mailing lists among students and staff, volunteered for the experiment. They were all naive to the purpose of the experiment, had normal or corrected-to-normal vision, and gave written and informed consent. The study conformed to the declaration of Helsinki, and was approved by the local ethical committee. Data from one participant was removed from the first experiment a posteriori because the tracking ratio was lower than 80% in the RE. Similarly, because of incorrect calibration of the eye-tracking device in the virtual conditions, data from 4 other participants was removed a posteriori from the first experiment, and of 3 participants from the second experiment. Therefore, only the data from sixteen participants (4F, 12M; age: avg.=24.9 ± 3.2, min=20, max=30) was used for the first experiment and the data from eighteen participants (4F, 14M; age: avg.=25.5 ± 4.0, min=20, max=36) was used for the second experiment.

#### 3.3 Analysis

##### 3.3.1 Eye Data Collection

It is important to distinguish the difference between gaze activity and eye movements. Eye movement refers to the local coordinates of the gaze relative to the head. Gaze activity, on the other hand, corresponds to the global coordinates of the gaze in the world space [20], which therefore also accounts for head rotations. In our case, we recorded eye movements, and assume that the head movements contribution is not significant. Qualitatively speaking we observe that participants do not significantly move the head when performing the task at hand, i.e. reaching destination 20m further down a 5m-wide street. The movement of the eyes recorded will therefore be close to the gaze activity, and this is the reason why we talk about eye-gaze activity in this work. Furthermore in case of sudden movements, it has been shown that the eyes initiate the movement, then the head, and finally the body, both in RE or VE [43, 47], thus resulting in a saccade. In this paper we are only interested by the location of fixations, and therefore will not analyse eye movements during saccades. In conclusion, to study eye-gaze activity we used the 2D location of the gaze in the recorded video of the environment seen by the participant. For the RE, a camera placed at the center of the eye-tracking glasses, just above the nose (see Figure 2-a), recorded what participants saw over time. Gaze coordinates correspond to pixel positions in this video. For the VE we recreated the same protocol by placing a camera with the same characteristics between the participant's two eyes in the VE (see Figure 2-b).

##### 3.3.2 Fixations Computation

Our eye-gaze activity analysis relies, as in many other studies on gaze, on the measure of visual fixations. As for other studies, our first important task is therefore to accurately register the fixations and the saccades [22]. Depending on the task and situation several methods have been proposed in the literature to compute fixations [26], each with advantages and limitations depending on the situation. In our situation, gaze fixations are computed based on the 2D gaze location of participants in the recorded images (real or virtual) over time.

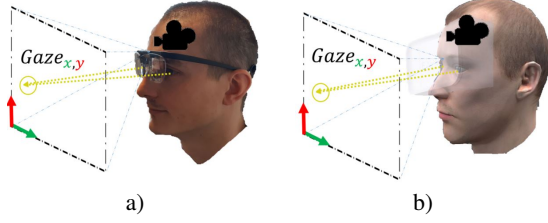


Figure 2: Setup used to collect gaze location for both real (a) and virtual (b) conditions. For each condition, 2D gaze location is displayed in the image recorded by the real or virtual camera (in black).

The method used to compute fixations is inspired by the dispersion-based algorithms (I-DT) described by Salvucci *et al.* [50]. We first compute the maximum distance between 2D gaze locations and their centroid over a sliding window of 100 ms. After identifying all the points where this distance is less than a threshold of  $1.5^\circ$ , we process each identified point by accumulating in the same fixation the neighbouring points if they respect these two conditions:

- $\text{mean}(\text{GazeD}) < 1.5^\circ$ ,
- $\text{std}(\text{GazeD}) - \text{StdInit} < 0.4^\circ$ ,

where  $\text{GazeD}$  is the list of distances between the fixation's centroid and all gaze points currently belonging to that fixation, and  $\text{StdInit}$  the standard deviation of  $\text{GazeD}$  at the start of the fixation. In our analyses, we chose a circle with a radius of  $1.5^\circ$ , as the diameter of the fovea is about  $3^\circ$ , while the threshold for variance was chosen empirically and set to  $0.4^\circ$ . Additional information about the pseudo-code for computing fixations is provided as supplemental material.

### 3.3.3 Independent Variables

As our goal is to understand eye-gaze activity while walking in crowds of different densities, as well as the biases that can be introduced by VR, we considered two main aspects of participants' gaze fixations. The first aspect relates to the characteristics of fixations, namely the average duration and the amplitude of saccades. The second aspect relates to where participants looked in the scene, through the coverage of the fixations:

#### Fixation descriptors

- *Average duration of fixations (ms)*. This feature informs about time spent on each object to extract visual information.
- *Amplitude of saccades (degree)*. This feature informs about the distance separating successive targets. It is computed as the distance between two successive fixations.

**eye-gaze spatial distribution** These features describe eye-gaze activity over the whole navigation task. For each condition and each participant, we computed a fixation map according to the definition given by Wooding *et al.* [60]. To create a fixation map, we start with a blank image and then, for each fixation, we add a gaussian at the center of the fixation. We use a standard deviation  $\sigma$  of  $1.5^\circ$  for the gaussian, which will approximate the fovea. The fixation map will be represented as a heat-map, and for the sake of visibility, we will display the logarithm values of this map. From this map, we calculated two metrics:

- $\text{PeakFixation}$  is the maximum value of the fixation map normalized over the number of fixations. It describes where participant's focused more their gaze during the task.

- *Coverage* the number of pixels in the fixation map superior to a threshold  $D_{crit}$  over the size of the image, as defined by Wooding *et al.* [60]. We choose  $D_{crit}$  so has to not consider isolated fixations, i.e., fixations at a larger distance than  $2\sigma$  ( $3^\circ$ ) from any another fixation. As a result,  $D_{crit}$  is computed as follows:

$$D_{crit} = \frac{1}{2\pi\sigma^2} + \frac{1}{2\pi\sigma^2}e^{-2} \quad (1)$$

where the first term of Equation 1 is the peak value of the gaussian representing each fixation in the fixation map, and the second term the value at  $2\sigma$ .

### 3.3.4 Statistical Analysis

We set the level of significance to  $\alpha = 0.05$ . A Shapiro Wilk test was performed to evaluate whether the distribution of our data followed a normal distribution. When comparing Real vs. Virtual conditions (Experiment 1), we conducted paired t-tests. In Experiment 2, we investigated the effect of density in Virtual conditions by conducting either a Friedman test with Wilcoxon-signed rank post tests when the distribution was not normal, and a one-way repeated measures analysis of variance (ANOVA) with post-hoc paired t-tests otherwise. Greenhouse-Geisser adjustments to the degrees of freedom were applied, when appropriate, to avoid any violation of the sphericity assumption. For the post-hoc tests, we adjusted the p value to account for multiple comparisons using the Benjamini-Hochberg procedure [4] with a false discovery rate of 0.1.

## 4 REAL VS. VIRTUAL VALIDATION

The goal of this first experiment is to evaluate whether the eye-gaze activity of a human walking a street is biased in VR conditions compared to real ones. While Berton *et al.* [6] showed qualitative similarities when considering an interaction between two walkers, it is not yet established how these results generalize to more complex scenarios involving larger numbers of pedestrians. In particular, our experiment aims at assessing the following hypotheses:

**H1.1** The scene is displayed through a HMD in the VE. This reduces the field of view of participants in comparison with RE. As a consequence, we expect gaze spatial distribution to be different in the VE. Consistently, we expect the amplitude of eye saccades, as well as the area covered by the gaze, to be smaller in the VE.

**H1.2** The feature of the RE are accurately reproduced in VE (same street, buildings, geometry and same density of people) and the participants' task remains identical. We therefore expect the duration of gaze fixations to be similar in both conditions, as participants should take similar visual information to perform the task in RE and VE.

**H1.3** The task is to walk toward the opposite side of the street, which is a central point in the participants field of vision. We thus expect to observe gaze fixations to be centered in the field of view.

### 4.1 Procedure

Participants were asked to physically walk through the real, and then, the virtual street (see Section 3). They performed the RE first because the parameters of the virtual condition were adjusted to be as similar as possible to the RE, in terms of visual density of pedestrians encountered. As all participants could not perform the RE under the exact same experimental conditions, we chose to minimize differences in terms of brightness and crowd by conducting the real condition during lunchtime over several days. The virtual counterpart was then conducted approximately one week later.



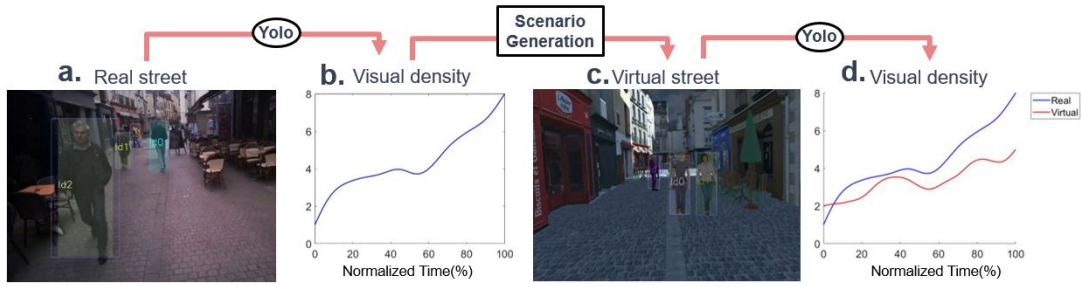


Figure 3: Generation of the virtual scenarios. a) Video recording when a participant walked in the real street. b) Density of people seen by participants over the normalized duration of the trial, estimated by tracking people visible in the video recording using deep learning algorithm. c) Virtual scenario, reproducing qualitatively similar situations in the virtual conditions in terms of virtual characters encountered, as seen by the participant. b) Density of individuals actually seen by the participant over the normalized duration in both the real (blue) and the virtual trial (red).

To enable comparisons between RE and VE, we first estimated for each trial the number of people they saw in the RE. We then created a specific stimuli that reproduced this same number. To this end, we detected and tracked people visible in the video recorded through the Tobii glasses (see Figure 3a-b) using a combination of two neural networks: Yolo [46] and DeepSort [58]. Based on tracking information, we categorized people into 3 categories: standing, walking in the same direction, or walking in the opposite direction to the participant. We generated scenarios with similar features by spawning virtual characters in the street accordingly (see Figure 3c-d for examples of generated feature values). The similarities between the generated RE and VE are analysed and discussed below.

Finally, participants wore Tobii eye-tracking glasses in the RE, whilst the Fove HMD was used in the VE. They performed 4 trials in each condition, as well as 2 initial training trials to get accommodated to VR. The experiment lasted approximately 10min for the RE, and approximately 15min for the VE.

## 4.2 Analysis & Results

### 4.2.1 Comparison between real and virtual stimuli

As mentioned above, the VE were generated so as to reproduce similar distributions of people compared to each corresponding real trial. These generated scenarios were also verified by the experimenter prior to the VE. To this end, we ran the same tracking techniques on virtual stimuli to estimate the number of characters seen by participants. Figure 4 presents the average number of individuals seen by each participant across trials, for both RE and VE. While inter-participant differences exist, and are expected as the RE could not be controlled in terms of pedestrian activity in the street, results show that the number of individuals seen is quite similar in both RE and VE, suggesting that the real and virtual stimuli presented were mostly similar in this aspect.

### 4.2.2 Fixations and Saccades

The average duration of fixations is illustrated in Figure 5-a) and is influenced by the condition ( $t(15)=3.9, p<0.005, r=0.71$ ), where the duration of fixations was significantly longer in VE ( $252.5 \pm 42.4ms$ ) than in RE ( $204.8 \pm 30.4ms$ ). The amplitude of saccades is illustrated in Figure 5-b) and is also influenced by the condition ( $t(15)=4.1, p<0.001, r=0.72$ ), where the amplitude of saccades is significantly larger in RE ( $6.9 \pm 1.3^\circ$ ) than in VE ( $5.3 \pm 1.4^\circ$ ).

### 4.2.3 Gaze spatial distribution

The average  $Peak_{Fixation}$  is illustrated in Figure 6-a), and is influenced by the condition ( $T = 0, Z = 3.51, p < 0.001, r = 0.88$ ), where results show that the peak value is significantly higher for the VE (0.18) than for the RE (0.11). Coverage is illustrated in

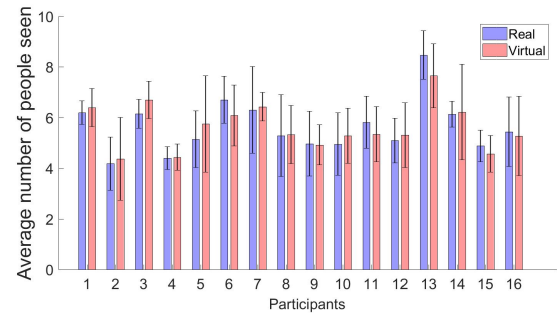


Figure 4: Average number of people seen by participants across trials for both RE and VE. The black bars represent the standard deviation for each participant and conditions.

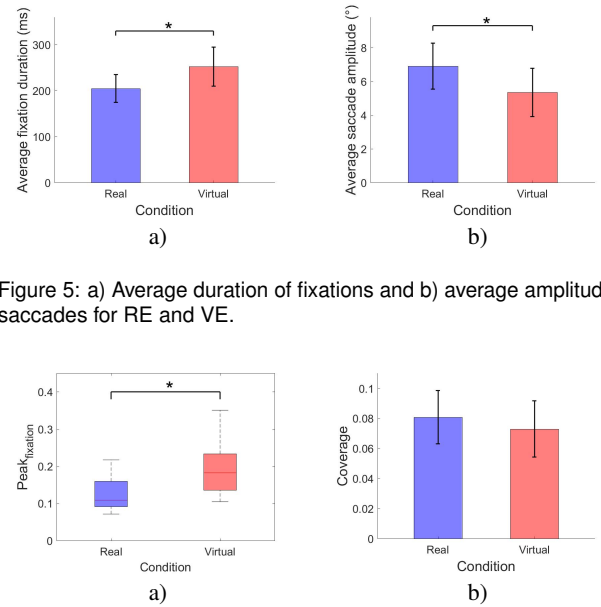


Figure 5: a) Average duration of fixations and b) average amplitude of saccades for RE and VE.

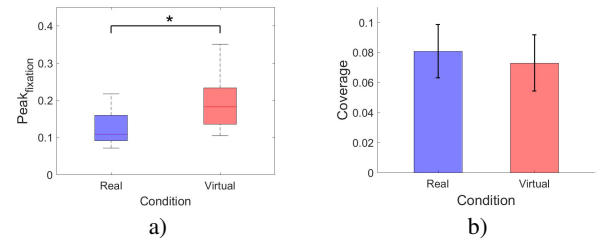


Figure 6: a)  $Peak_{Fixation}$  and b) Coverage for RE and VE.

Figure 6-b), but we did not find significant differences between RE and VE ( $p = 0.14$ ).

A visual representation of these results is provided in Figure 7, which displays the fixation maps for both the RE and VE. In particular, it is noticeable that gaze locations seem to follow a centered

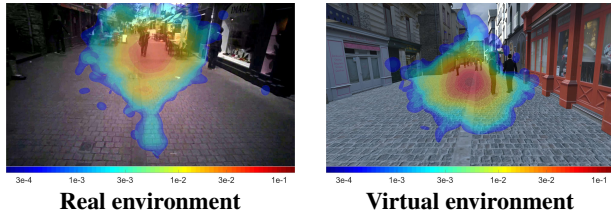


Figure 7: Heat-map (log-transformed) of the gaze fixation distribution for both RE and VE

distribution for both conditions. While the center of this distribution is approximately at the center of the  $1920 \times 1080$  image on the x-axis for both conditions (RE:  $Peak_x = 965 \pm 56pixel$ , VE:  $Peak_x = 962 \pm 68pixel$ ), it appears to be higher on the y-axis for the RE than for the real one (RE:  $Peak_y = 717 \pm 102pixel$ , VE:  $Peak_y = 526 \pm 65pixel$ ). To identify whether this difference could be due to a shift of the horizontal reference axis between the RE and VE (i.e., an angle between the Tobii camera axis and HMD virtual camera axis), we asked two volunteers to identify (by clicking with the mouse button) the horizon line at the end of the street in a selection of 100 images randomly extracted from the video recordings of real trials, as well as in 100 images from virtual trials. The distribution of the altitude of the horizon line in images follows a normal distribution with a average coordinate of  $804.5 \pm 112.8 pixel$  for the real images and an average coordinate of  $597.6 \pm 86.6 pixel$  for the virtual ones, showing therefore a difference of 206.9 pixels between these two centres (equivalent to an angular difference of  $9.7^\circ$ ).

#### 4.2.4 Gaussian model for gaze prediction

The spatial distribution of the fixations follows a centered distribution that decreases exponentially around the center, as shown in the log scale heat-maps in Figure 7. We fit a Gaussian model on this distribution, which estimated parameters ( $\mu_x$ ,  $\mu_y$ ,  $\sigma_x$  and  $\sigma_y$ ) are presented in Table 1 with the corresponding coefficient of correlation ( $R^2$ , with  $p < 0.05$ ). It is also interesting to mention that the vertical difference in the location of the center of gaze fixations as estimated by our Gaussian model fitting is of  $7.75^\circ$ , which is relatively close to the angular difference observed in the altitude of horizon estimated from the real and virtual images ( $9.7^\circ$ ).

### 4.3 Discussion

In this first experiment we show that there are several differences in the walker’s eye-gaze activity between the RE and VE. First, saccades are of a greater amplitude in the RE than in the VE. In addition, the value of  $Peak_{fixation}$  is different, indicating that the gaze is more intensely focused on the center of the visual field in VR. Moreover, the fitting of the Gaussian model on the distribution of fixations location distribution resulted in a correlation coefficient of 0.77 for both conditions, that we estimate to be sufficiently high to validate this choice. A larger value is computed for  $\sigma$  in the RE. All these differences are consistent with hypothesis **H1.1**. Nevertheless, no statistical difference is observed for coverage, which does not allow us to fully confirm the hypothesis. Note that our hypothesis was founded on the restriction of the field of view due to the use of HMDs. However, the performed task, i.e., walking straight in a street, may not require the exploration of peripheral areas. If true, users may not have to compensate such a reduction.

A statistical difference is however present for the duration of fixation, thus invalidating our second hypothesis **H1.2**. We interpret that the cause of this difference is possibly due to the level of detail in the digital replica of the street. In spite of the high quality of the digital scene, there are always missing details, e.g., birds, or sounds,

Conditions	$\mu_x$	$\mu_y$	$\sigma_x^\circ$	$\sigma_y^\circ$	$R^2$
RE	0.15°	8.10°	7.69°	6.93°	0.77
VE	-0.57°	0.35°	6.11°	5.08°	0.77

Table 1:  $\mu_x$ ,  $\mu_y$ ,  $\sigma_x$  and  $\sigma_y$  of the gaussian distribution for the gaze location in the image and  $R^2$  between this distribution and the initial distribution with respect to the experimental conditions.

that could have attracted the visual attention of our participants in RE, and provoked faster fixations.

We also find a significant difference in the altitude of the gaze center between RE and VE. This difference appears to be well explained by a difference in the orientation of the eye-trackers reference horizontal axis between the two conditions. Finally, participants display a similar eye-gaze behavior as they have a centered distribution for gaze location toward their goal, validating thus Hypothesis **H1.3**.

Therefore, our experiment highlights the importance of the digital content when performing gaze studies in VR, as eye-gaze activity seems to be affected by the richness and realism of this content. Despite our experiment being based on high-fidelity urban digital mockups, we cannot claim that we were able to reproduce all the details of reality. Nevertheless, these results provide valuable insights to understand which aspects of eye-gaze activity are qualitatively similar between RE and VE, to further explore using VR aspects about navigating in a crowd which are not possible in real situations. We thus believe that VR is a valid tool to study such activity, while experimenters should remain aware that some elements in real environments may distract walkers’ gaze more.

## 5 EFFECT OF CROWD DENSITY ON EYE-GAZE BEHAVIOUR

The goal of this experiment is to evaluate the impact of crowd density on eye-gaze activity, by leveraging the use of VR to accurately control the number of displayed characters. When density increases, walkers have more frequent interactions and face more people. In this experiment we aim at assessing the following hypotheses:

**H2.1** As the amount of visual information increases with density, we expect that the duration of fixations will decrease accordingly. By increasing the scanning frequency, participants will be able to consider more elements in the VE.

**H2.2** In addition, we also expect participants’ gaze to be more focused towards the center of the visual field, since pedestrians with the highest risk of collision are typically the ones in front of them. In particular, we expect that the amplitude of the saccades, as well as the area covered by the gaze, will decrease as the density of the crowd increases. This also implies that the gaze will be more intense in the center of the field of vision.

### 5.1 Procedure

In Experiment 2, participants were immersed in the same VE as described in Section 3, and were instructed to navigate in the virtual street through a crowd walking in a unidirectional flow in the opposite direction. Our objective for this experiment is to study the impact of crowd density on eye-gaze activity. In particular, for a specific crowd density  $d$ , we generated a scenario with  $d$  virtual characters every 15 meters at the start of the trial. These characters were driven by the RVO crowd simulator, similarly to the previous experiment (see details in Section 3.1). Participants were asked to navigate through the crowd for approximately 20m, so different speeds of virtual humans did not affect the visual densities for such a short period. Participants were presented in random order with 6 different conditions of densities:  $d \in [2, 5, 10, 14, 18, 24]$  (Figure 10). For each density, participants had to perform 4 repetitions. Participants performed in total 24 trials, and the experiment took on average 20 minutes per participant.

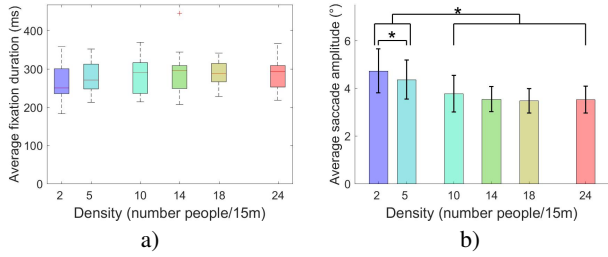


Figure 8: a) Average duration of fixations and b) average participant's amplitude of saccades depending on crowd density.

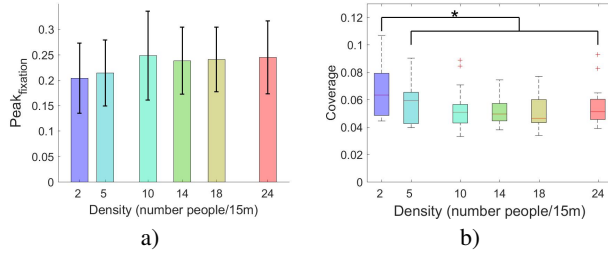


Figure 9: a)  $Peak_{Fixation}$  and b) Coverage depending on crowd density.

Table 2:  $\mu_x$ ,  $\mu_y$ ,  $\sigma_x$  and  $\sigma_y$  of the gaussian distribution for the gaze location in the image and  $R^2$  between this distribution and the initial distribution with respect to the density conditions.

Conditions	$\mu_x$	$\mu_y$	$\sigma_x^\circ$	$\sigma_y^\circ$	$R^2$
$d : 2$	$-0.72^\circ$	$0.02^\circ$	$5.06^\circ$	$5.82^\circ$	0.75
$d : 5$	$-1.08^\circ$	$-0.11^\circ$	$4.38^\circ$	$5.39^\circ$	0.79
$d : 10$	$-1.22^\circ$	$-0.55^\circ$	$4.21^\circ$	$4.64^\circ$	0.83
$d : 14$	$-0.69^\circ$	$-1.01^\circ$	$4.26^\circ$	$4.53^\circ$	0.84
$d : 18$	$-1.05^\circ$	$-0.91^\circ$	$4.07^\circ$	$4.01^\circ$	0.87
$d : 24$	$-0.87^\circ$	$-1.36^\circ$	$4.26^\circ$	$4.45^\circ$	0.83

## 5.2 Analysis & Results

### 5.2.1 Fixation descriptors

The analysis on average duration of fixations shows an effect of density ( $\chi^2(5) = 14.15873$ ,  $p < 0.01463$ ). However, this is not confirmed by post-hoc pairwise comparisons. Nevertheless, for illustrative purposes, this result is displayed in Figure 8-a. Density has however a strong effect on the amplitude of saccades ( $F(2.69, 45.8) = 18.4$ ,  $p < 0.000001$ ,  $\eta_p^2 = 0.51$ ), where post-hoc analysis shows that the amplitude of saccades is significantly larger when navigating in a crowd with a density of 2 and 5 than for any other condition (Figure 8-b).

### 5.2.2 Gaze spatial Distribution

The average  $Peak_{Fixation}$  is illustrated in Figure 9-a). An ANOVA shows an effect of the density ( $F(5, 85) = 2.74$ ,  $p < 0.05 (= 0.024)$ ,  $\eta_p^2 = 0.14$ ), which is not confirmed by post-hoc pairwise comparisons. We however find an effect of density on Coverage ( $\chi^2(5) = 15.77778$ ,  $p = 0.00751$ ), where post-hoc analysis shows that the coverage is larger when navigating in a crowd with a density of 2 than for any other condition (Figure 9-b).

Fixation maps are also displayed for each condition in Figure 10. For each density, the gaze location follows a centered distribution, furthermore it seems that the coverage by the gaze is decreasing as the density of the crowd is increasing, especially on the vertical-axis.

### 5.2.3 Gaussian model for gaze prediction

As in the precedent experiment, the spatial distribution of the fixations follows a centered distribution that decreases around the center,

as shown in the log scale heat-maps in Figure 10. We fitted a Gaussian model on the gaze location distribution. The estimated Gaussian parameters ( $\mu_x$ ,  $\mu_y$ ,  $\sigma_x$  and  $\sigma_y$ ) are reported in Table 2, with the corresponding correlation coefficient ( $R^2$ , with  $p < 0.05$ ). Qualitatively, the  $\sigma$  values decreased with the increase of density, especially on the vertical-axis.

## 5.3 Discussion

In this second experiment we show some differences in the walker's eye-gaze activity when crowd density increases. In particular, we show statistical differences in the amplitude of the saccades and the area covered by the gaze. These two variables increase when the crowd density decreases. Furthermore, this result is also supported by the Gaussian model fittings. The  $\sigma$  parameter, which is correlated with the coverage, decreases with the increasing crowd density, especially on vertical-axis. However, there is no statistical difference for  $Peak_{Fixation}$ , which indicates that walker's gaze is not more intense in the center of the field of vision as expected. From all these results, we can only partially validate our second hypothesis **H2.2**. Nevertheless, it seems that a Gaussian model could be used to estimate the distribution of gaze location as the correlation coefficient is high for each condition of density.

Concerning the duration of the fixations, post-hoc analysis did not allow to validate that density has an impact on this data, thus invalidating our hypothesis **H2.1**. Participants took the same amount of time to search for visual information regardless of the crowd density present in front of them. Finally, it can be noted that regardless of crowd density, participants display mostly a similar eye-gaze behaviour, as they show a centered distribution of gaze positions toward their objective. These results are in line with previous work [61] showing that eye-gaze activity is dependent of the task. In conclusion, these results indicate that participants' gaze is more focused toward the direction of the goal they have to reach as crowd density increases. Their scanning range decreases, showing that they visually take into account mainly pedestrians and visual cues in front of them.

## 6 GENERAL DISCUSSION

### 6.1 Crowd simulation

The motivation of this work was the modelling and simulation of crowds behaviour. In particular, Experiment 2 presented in Section 5 explores the question of interaction neighborhood, and how it varies with increasing densities. We first show that density has no significant effect on the duration of fixations. From the perspective of crowd modelling, this let us think that the number of interaction neighbors tends to remain constant with density. Indeed, an increase in the number of neighbors would probably have asked participants to visually scan them faster. For a similar situation such as an opposite crowd in a street, we thus recommend to work with a constant number of neighbors in simulations (i.e., the number of simulated interactions for each agent). A more detailed analysis of the observed characters would however be required to evaluate whether some of them are observed multiple times.

Concerning the selection features to use to select neighbor agents, we recommend ordering them by risk of collision (which can be estimated in different ways, as discussed in [36]). Indeed, even though coverage did not significantly change with density, our results reveal that gaze tends to refocus around the visual center when density increases: this is the area where characters presenting the highest risk of collision in this bidirectional traffic condition.

### 6.2 Limitations

Section 4.3 attributes changes in eye-gaze behaviour between RE and VE to the lack of some details in the virtual scene. This interpretation is corroborated by previous studies [53]. In particular, the digital scene did not incorporate sound simulation. Locomotion and



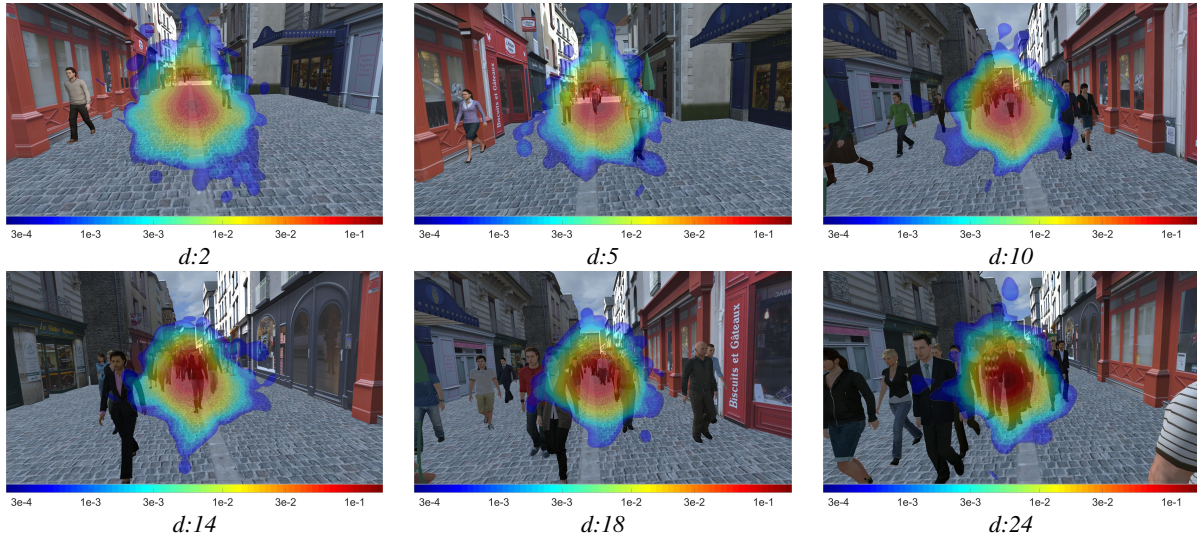


Figure 10: Virtual street with all the different crowd densities. For each density, the fixation area (log-transformed) is displayed on top of the image

collision avoidance, is mainly controlled from visual information. As an example, Silva [52] showed that, with 3 pedestrians or more, the emission of sound has no effect on the manoeuvres performed by participants. However, gaze activity was not explored in this work, and a sound can certainly easily attract our attention, thus impacting some characteristics of our gaze activity [14].

In the second experiment on crowd density, virtual humans had to avoid any collision with the participant. We did not want to have participants traversed by virtual characters, that would have negatively impacted immersion. In RE, the level of attention paid by surrounding people to their navigation can vary a lot. As an example, the use of cellular phone while walking affect gait kinematics when avoiding other pedestrians [30]. These differences in the behaviour of neighbors may also induce a change in the participants' behaviour. We believe that, generally, the reduction of the field of view by the Fove HMD may have an impact on the eye-gaze behaviour. This was already outlined by previous studies [6, 44]. However, in our case, we think that this impact was limited, especially because the task was to walk straight ahead in a street. The goal as well as the oncoming obstacles were always visible in the central vision area. Nevertheless, we are interested in extending the number of situations and to address the case of crossing traffic. The width of the field of view would certainly take a greater importance then. It is also possible that performing the task in a street (RE and VE) may have overly normalized eye-gaze behaviour, leading to little effect of density on the studies variables. In a next step, we would therefore be interested in studying such eye-gaze behaviours in more open places, in order to evaluate whether normalizing the reaction of our participants in a closed-space (street) was indeed overly constraining their eye-gaze activity. In addition, it is important to note that the method we used to compute fixations cannot be applied to all situations, as it assumes small head movements. We used this method for a fair comparison with RE where the capture of head movement in a street is challenging. With the use of VR, we would then be able to compute these motions in order to adapt our analysis to more complex scenario. Furthermore, several recent studies explored the coupled analysis of locomotion and gaze and considered, for instance, walking speed [18]. In this work we have focused solely on the eye-gaze activity, but we intend to focus on this type of study in our future work. Finally, it would have been interesting to have a larger number of participants in this study, so as to improve the accuracy of the fitted Gaussian models as well as to study more

sensitive metrics, such as the inter-individual variability in gaze data in VR [19], which we plan to explore in the future.

## 7 CONCLUSION & FUTURE WORKS

In this paper we have carried out two experiments on the study of eye-gaze activity while walking in a street. Our first experiment was to study the activity of the gaze while navigating in a real and a virtual environment in order to evaluate the impact of VR on the eye-gaze activity. Our results show a qualitatively similar eye-gaze activity with some quantitative differences. In our second experiment we studied the impact of crowd density on a walker's gaze in VR. Our results demonstrate an influence of density on gaze deployment, where it decreases as the density increases while navigating in a street full with an opposite crowd. For such situation, this indicates that in high densities, walkers have a tendency to focus more their gaze in front of them. Consequently, they will visually take more into account people in front them than people in their surroundings. We are able to provide guidelines for the design of models of local interaction for crowd simulators.

More importantly, this work opens new perspectives for future research. The modeling of crowds raises plenty of questions. In first place, we have studied the effect of density in the case of an oncoming traffic in a street only. It is first required to explore more traffic conditions to get a deeper understanding of interaction neighborhood. For example, in the case of crossing flows or in a more open space, we expect the gaze to explore more the peripheral areas of the field of vision. It would be also very interesting to try to saturate one's visual field with many interactions of importance (e.g., many characters all on a collision course) to explore the limits of visual integration and observe if walkers apply a specific strategy in such cases. Nevertheless, if peripheral vision gets more important in new scenarios, it would be certainly required to re-evaluate eye-gaze activity when using HMDs with a wide field of vision, that are becoming more and more popular. Finally, as we have highlighted the possible role of other sensory channels on visual attention (e.g., sound or touch), we would like to integrate higher fidelity scenes and VR rendering techniques in our experimental VR platform.

## ACKNOWLEDGMENTS

This work was funded by the ANR OPMoPS project (ANR-16-SEBM-0004) and the Inria Associate Team BEAR.

## REFERENCES

- [1] P. Agethen, V. S. Sekar, F. Gaisbauer, T. Pfeiffer, M. Otto, and E. Rukzio. Behavior analysis of human locomotion in the real world and virtual reality for the manufacturing industry. *ACM Transactions on Applied Perception (TAP)*, 15(3):20, 2018.
- [2] F. Argelaguet Sanz, A.-H. Olivier, G. Bruder, J. Pettré, and A. Lécuyer. Virtual proxemics: Locomotion in the presence of obstacles in large immersive projection environments. In *2015 IEEE Virtual Reality (VR)*, pages 75–80. IEEE, 2015.
- [3] P. Basili, M. Sağlam, T. Kruse, M. Huber, A. Kirsch, and S. Glasauer. Strategies of locomotor collision avoidance. *Gait & posture*, 37(3):385–390, 2013.
- [4] Y. Benjamini and Y. Hochberg. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal statistical society: series B (Methodological)*, 57(1):289–300, 1995.
- [5] D. Bernardin, H. Kadone, D. Bennequin, T. Sugar, M. Zaoui, and A. Berthoz. Gaze anticipation during human locomotion. *Experimental brain research*, 223(1):65–78, 2012.
- [6] F. Berton, A.-H. Olivier, J. Bruneau, L. Hoyet, and J. Pettré. Studying gaze behaviour during collision avoidance with a virtual walker: Influence of the virtual reality setup. 2019.
- [7] H. Brument, L. Podkosova, H. Kaufmann, A. H. Olivier, and F. Argelaguet. Virtual vs. physical navigation in vr: Study of gaze and body segments temporal reorientation behaviour. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pages 680–689. IEEE, 2019.
- [8] J. Bruneau, A.-H. Olivier, and J. Pettré. Going through, going around: A study on individual avoidance of groups. *IEEE transactions on visualization and computer graphics*, 21(4):520–528, 2015.
- [9] M. A. Bühler and A. Lamontagne. Circumvention of pedestrians while walking in virtual and physical environments. *IEEE transactions on neural systems and rehabilitation engineering*, 26(9):1813–1822, 2018.
- [10] M. A. Bühler and A. Lamontagne. Circumvention of pedestrians while walking in virtual and physical environments. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 2018.
- [11] F. Capozzi, A. P. Bayliss, and J. Ristic. Gaze following in multiagent contexts: Evidence for a quorum-like principle. *Psychonomic bulletin & review*, 25(6):2260–2266, 2018.
- [12] U. Chattaraj, A. Seyfried, and P. Chakroborty. Comparison of pedestrian fundamental diagram across cultures. *Advances in complex systems*, 12(03):393–405, 2009.
- [13] M. E. Cinelli, A. E. Patla, and F. Allard. Behaviour and gaze analyses during a goal-directed locomotor task. *The Quarterly Journal of Experimental Psychology*, 62(3):483–499, 2009.
- [14] A. Coutrot, N. Guyader, G. Ionescu, and A. Caplier. Influence of soundtrack on eye movements during video exploration. 2012.
- [15] J. L. Croft and D. Panchuk. Watch where you’re going? interferer velocity and visual behavior predicts avoidance strategy during pedestrian encounters. *Journal of Motor Behavior*, 0(0):1–11, 2017.
- [16] P. Dickinson, K. Gerling, K. Hicks, J. Murray, J. Shearer, and J. Greenwood. Virtual reality crowd simulation: effects of agent density on user experience and behaviour. *Virtual Reality*, 23(1):19–32, 2019.
- [17] M. Dicks, C. Clashing, L. O’Reilly, and C. Mills. Perceptual-motor behaviour during a simulated pedestrian crossing. *Gait & posture*, 49:241–245, 2016.
- [18] H. Dietrich and M. Wuehr. Strategies for gaze stabilization critically depend on locomotor speed. *Neuroscience*, 408:418–429, 2019.
- [19] M. Dorr, T. Martinetz, K. R. Gegenfurtner, and E. Barth. Variability of eye movements when viewing dynamic natural scenes. *Journal of vision*, 10(10):28–28, 2010.
- [20] D. Guitton and M. Volle. Gaze control in humans: eye-head coordination during orienting movements to targets within and beyond the oculomotor range. *Journal of neurophysiology*, 58(3):427–459, 1987.
- [21] D. Helbing and P. Molnar. Social force model for pedestrian dynamics. *Physical review E*, 51(5):4282, 1995.
- [22] R. S. Hessels, D. C. Niehorster, M. Nyström, R. Andersson, and I. T. Hooge. Is the eye-movement field confused about fixations and saccades? a survey among 124 researchers. *Royal Society open science*, 5(8):180502, 2018.
- [23] M. Huber, Y.-H. Su, M. Krüger, K. Faschian, S. Glasauer, and J. Hermsdörfer. Adjustments of Speed and Path when Avoiding Collisions with Another Pedestrian. *PLoS ONE*, 9(2):e89589, 2014.
- [24] J. Jovancevic, B. Sullivan, and M. Hayhoe. Control of attention and gaze in complex environments. *Journal of Vision*, 6(12):9–9, 2006.
- [25] J. Jovancevic-Misic and M. Hayhoe. Adaptive gaze control in natural environments. *Journal of Neuroscience*, 29(19):6234–6238, 2009.
- [26] A. Kar and P. Corcoran. A review and analysis of eye-gaze estimation systems, algorithms and performance evaluation methods in consumer platforms. *IEEE Access*, 5:16495–16519, 2017.
- [27] K. Kitazawa and T. Fujiyama. Pedestrian vision and collision avoidance behavior: Investigation of the information process space of pedestrians using an eye tracker. In *Pedestrian and evacuation dynamics 2008*, pages 95–108. Springer, 2010.
- [28] A. Knorr, L. Willacker, J. Hermsdörfer, S. Glasauer, and M. Krüger. Influence of Person- and Situation-Specific Characteristics on Collision Avoidance Behavior in Human Locomotion. *J. of Exp. Psy.: Human Perception and Performance*, 42(9):1332–1343, 2016.
- [29] M.-L. Lai, M.-J. Tsai, F.-Y. Yang, C.-Y. Hsu, T.-C. Liu, S. W.-Y. Lee, M.-H. Lee, G.-L. Chiou, J.-C. Liang, and C.-C. Tsai. A review of using eye-tracking technology in exploring learning from 2000 to 2012. *Educational research review*, 10:90–115, 2013.
- [30] S. Licence, R. Smith, M. P. McGuigan, and C. P. Earnest. Gait pattern alterations during walking, texting and walking and texting during cognitively distractive tasks while negotiating common pedestrian obstacles. *PLoS one*, 10(7):e0133281, 2015.
- [31] J. M. Loomis, J. M. Knapp, et al. Visual perception of egocentric distance in real and virtual environments. *Virtual and adaptive environments*, 11:21–46, 2003.
- [32] S. Lynch, J. Pettré, J. Bruneau, R. Kulpa, A. Cretual, and A.-H. Olivier. Effect of virtual human gaze behaviour during an orthogonal collision avoidance walking task. In *IEEE Virtual Reality*, 2018.
- [33] S. D. Lynch, R. Kulpa, L. A. Meerhoff, J. Pettré, A. Cretual, and A. H. Olivier. Collision avoidance behavior between walkers: global and local motion cues. *IEEE Transactions on Visualization and Computer Graphics*, PP(99):1–1, 2017.
- [34] D. S. Marigold and A. E. Patla. Gaze fixation patterns for negotiating complex ground terrain. *Neuroscience*, 144(1):302–313, 2007.
- [35] S. A. McMains and S. Kastner. Visual attention. *Encyclopedia of neuroscience*, pages 4296–4302, 2009.
- [36] L. Meerhoff, J. Bruneau, A. Vu, A.-H. Olivier, and J. Pettré. Guided by gaze: Prioritization strategy when navigating through a virtual crowd can be assessed through gaze activity. *Acta psychologica*, 190:248–257, 2018.
- [37] L. A. Meerhoff, J. Pettré, S. D. Lynch, A. Crétual, and A.-H. Olivier. Collision avoidance with multiple walkers: Sequential or simultaneous interactions? *Frontiers in psychology*, 9:2354, 2018.
- [38] C. Mousas, A. Koiliak, D. Anastasiou, B. Rekadbar, and C.-N. Anagnostopoulos. Effects of self-avatar and gaze on avoidance movement behavior. 2019.
- [39] A. H. Olivier, J. Bruneau, R. Kulpa, and J. Pettré. Walking with virtual people: Evaluation of locomotion interfaces in dynamic environments. *IEEE Transactions on Visualization and Computer Graphics*, PP(99):1–1, 2017.
- [40] A.-H. Olivier, A. Marin, A. Crétual, A. Berthoz, and J. Pettré. Collision avoidance between two walkers: Role-dependent strategies. *Gait & Posture*, 38(4):751–756, 2013.
- [41] A.-H. Olivier, A. Marin, A. Crétual, and J. Pettré. Minimal predicted distance: A common metric for collision avoidance during pairwise interactions between walkers. *Gait and Posture*, 36(3):399–404, 2012.
- [42] R. Parasuraman and M. Rizzo. *Neuroergonomics: The brain at work*. Oxford University Press, 2008.
- [43] A. E. Patla. Understanding the roles of vision in the control of human locomotion. *Gait & Posture*, 5(1):54–69, 1997.
- [44] K. Pfeil, E. M. Taranta II, A. Kulshreshtha, P. Wisniewski, and J. J. LaViola Jr. A comparison of eye-head coordination between virtual and physical realities. In *Proceedings of the 15th ACM Symposium on Applied Perception*, page 18. ACM, 2018.
- [45] J. M. Plumert, J. K. Kearney, J. F. Cremer, and K. Recker. Distance

- perception in real and virtual environments. *ACM Transactions on Applied Perception (TAP)*, 2(3):216–233, 2005.
- [46] J. Redmon and A. Farhadi. Yolov3: An incremental improvement. *arXiv*, 2018.
  - [47] R. Reed-Jones, J. Reed-Jones, L. A. Vallis, and M. Hollands. The effects of constraining eye movements on visually evoked steering responses during walking in a virtual environment. *Experimental brain research*, 197(4):357–367, 2009.
  - [48] K. W. Rio, G. C. Dachner, and W. H. Warren. Local interactions underlying collective motion in human crowds. *Proceedings of the Royal Society B: Biological Sciences*, 285(1878):20180611, 2018.
  - [49] M. Rubo and M. Gamer. Virtual reality as a proxy for real-life social attention? In *Proceedings of the 2018 ACM Symposium on Eye Tracking Research & Applications*, page 81. ACM, 2018.
  - [50] D. D. Salvucci and J. H. Goldberg. Identifying fixations and saccades in eye-tracking protocols. In *Proceedings of the 2000 symposium on Eye tracking research & applications*, pages 71–78. ACM, 2000.
  - [51] A. Seyfried, B. Steffen, W. Klingsch, and M. Boltes. The fundamental diagram of pedestrian movement revisited. *Journal of Statistical Mechanics: Theory and Experiment*, 2005(10):P10002, 2005.
  - [52] W. S. Silva, G. Aravind, S. Sangani, and A. Lamontagne. Healthy young adults implement distinctive avoidance strategies while walking and circumventing virtual human vs. non-human obstacles in a virtual environment. *Gait & posture*, 61:294–300, 2018.
  - [53] M. Slater, P. Khanna, J. Mortensen, and I. Yu. Visual realism enhances realistic response in an immersive virtual environment. *IEEE computer graphics and applications*, 29(3):76–84, 2009.
  - [54] J. Snape, S. J. Guy, D. Vembar, A. Lake, M. C. Lin, and D. Manocha. Reciprocal collision avoidance and navigation for video games. In *Game Developers Conf., San Francisco*, 2012.
  - [55] J. Van Den Berg, S. Patil, J. Sewall, D. Manocha, and M. Lin. Interactive navigation of multiple agents in crowded environments. In *Symp. on Interactive 3D graphics and games*, pages 139–147, 2008.
  - [56] N. Wang, S. Xu, S. Zhang, Y. Luo, and H. Geng. Erp evidence on how gaze convergence affects social attention. *Scientific reports*, 9(1):7586, 2019.
  - [57] W. H. Warren. Collective motion in human crowds. *Current directions in psychological science*, 27(4):232–240, 2018.
  - [58] N. Wojke, A. Bewley, and D. Paulus. Simple online and realtime tracking with a deep association metric. In *2017 IEEE International Conference on Image Processing (ICIP)*, pages 3645–3649. IEEE, 2017.
  - [59] D. Wolinski, M. C. Lin, and J. Pettré. Warpdriever: context-aware probabilistic motion prediction for crowd simulation. *ACM Transactions on Graphics (TOG)*, 35(6):164, 2016.
  - [60] D. S. Wooding. Eye movements of large populations: Ii. deriving regions of interest, coverage, and similarity using fixation maps. *Behavior Research Methods, Instruments, & Computers*, 34(4):518–528, 2002.
  - [61] A. L. Yarbus. *Eye movements and vision*. Springer, 2013.
  - [62] H. Yeh, S. Curtis, S. Patil, J. van den Berg, D. Manocha, and M. Lin. Composite agents. In *Proceedings of the 2008 ACM SIGGRAPH/Eurographics Symposium on Computer Animation*, pages 39–47. Eurographics Association, 2008.