Modeling and validation of dynamics of social influence in virtual human crowds

Elham Mohammadi Jorjafki
Humans are social by nature, and therefore when they come together in a large group, the resulting crowd behavior is difficult to explain in terms of individual motion dynamics. In pedestrian crowds, social influence occurs in the form of collision avoidance and in following neighbors’ actions. Understanding the dynamics of social influence can help improve crowd management strategies. Obtaining experimental data with crowds, however, is impractical due to the number of subjects needed and unsafe due to possible hazards of having many people assemble in a single place. The goal of this research is to develop a realistic virtual reality platform for crowd behavior studies and use it to validate the effect of social influence in human crowds. The specific objective of this research is to use methods from dynamical systems modeling to enable a custom virtual environment of an interactive humanoid crowd. The virtual environment is then evaluated through a behavioral experiment. Specifically, in this thesis we first adapt pedestrian motion models from the literature to produce dense pedestrian crowds where individuals can maintain close proximity and avoid unrealistic turning motions. We modify the collision avoidance and interaction functions so that the computational burden is reduced to realistically simulate human crowds of up to
sixty individuals at real-time speeds on a 3D graphics environment. The crowd simulation is then utilized to test the hypothesis that behavioral contagion triggered in the form of a subset of virtual characters shifting their gaze upwards can induce a human participant to follow that action. A human-subjects study with seventy participants is conducted to assess the performance of the virtual environment as well as test the hypothesis. Participants are surveyed on the believability of the virtual environment and their postural responses measured to record the effect of behavioral contagion.

Our results show that the virtual environment is effective in being perceived as natural and realistic and that participants are able to interact with the virtual crowd without experiencing any visual delays. We find that participants follow the visual gaze of virtual characters and that this action is proportional to the number of characters that trigger such a gaze. Contagion effect is further established in terms of the time spent looking up by the participants which is also found to be dependent on the size of the stimulus group.
MODELING AND VALIDATION OF DYNAMICS OF SOCIAL INFLUENCE
IN VIRTUAL HUMAN CROWDS

BY
ELHAM MOHAMMADI JORJAFKI
© 2018 Elham Mohammadi Jorjafki

A THESIS SUBMITTED TO THE GRADUATE SCHOOL
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE
MASTER OF SCIENCE

DEPARTMENT OF MECHANICAL ENGINEERING

Thesis Director:
Sachit Butail
ACKNOWLEDGEMENTS

I would like to thank my thesis advisor, Dr. Sachit Butail, who helped and supported me through all the challenges. He was always ready for help and collaboration in all steps of the research. I appreciate the time, energy and knowledge that he invested in my research.

In addition, I would like to thank the other thesis committee members, Dr. Brad Sagarin and Dr. Ji-Chul Ryu, for their collaboration and insightful comments.

I would also like to express my gratitude to Dr. Michael E. Papka, who kindly invited us to his lab and provided an opportunity for us to experience virtual reality platforms and broaden our knowledge in this area in the early steps of the research.

My sincere thanks also goes to Edouard Brooks, who helped me a lot with the challenges I faced in the phone-based virtual reality development. He generously shared his previous experience with me and even came to our lab, which I am more than grateful for.

I also want to thank my fellow lab members in EASeL: Kiran Maridi, Joseph Kempel, Hari Boddeeti, Kate Chwistek and Amro Shanshal. They helped us with behavior encoding and were always supportive. In addition, I thank all the students and Omron lab managers for their cooperation during experiment conduction.

Finally, I wanted to express my thoughtful gratitude to my parents, my brothers, and my partner, who always backed me with their unconditional love and encouragement.
DEDICATION

To my parents Reza and Batool and my beloved brothers Amin and Moein
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>LIST OF TABLES</th>
<th>vi</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIST OF FIGURES.</td>
<td>vii</td>
</tr>
<tr>
<td>Chapter</td>
<td></td>
</tr>
<tr>
<td>1 INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>1.1 Background</td>
<td>3</td>
</tr>
<tr>
<td>1.2 Thesis Contribution</td>
<td>4</td>
</tr>
<tr>
<td>2 CROWD BEHAVIOR MODELS</td>
<td>6</td>
</tr>
<tr>
<td>2.1 Pedestrian Dynamics</td>
<td>6</td>
</tr>
<tr>
<td>2.2 Social Force Model</td>
<td>7</td>
</tr>
<tr>
<td>2.3 Anticipatory Collision Avoidance Model and Variants</td>
<td>8</td>
</tr>
<tr>
<td>3 VIRTUAL REALITY ENVIRONMENT</td>
<td>11</td>
</tr>
<tr>
<td>3.1 Virtual Environment Design</td>
<td>11</td>
</tr>
<tr>
<td>3.2 Virtual Crowd Simulation</td>
<td>14</td>
</tr>
<tr>
<td>3.3 Virtual Crowd Trajectories Within the Indoor Lab</td>
<td>15</td>
</tr>
<tr>
<td>3.4 Virtual Reality Platform</td>
<td>16</td>
</tr>
<tr>
<td>3.4.1 Samsung Gear VR</td>
<td>17</td>
</tr>
<tr>
<td>3.4.2 HTC Vive</td>
<td>18</td>
</tr>
<tr>
<td>3.5 Real-Time Performance</td>
<td>20</td>
</tr>
<tr>
<td>4 HUMAN-SUBJECTS STUDY ON BEHAVIORAL CONTAGION</td>
<td>25</td>
</tr>
<tr>
<td>4.1 Behavioral Contagion Experimental Study</td>
<td>25</td>
</tr>
<tr>
<td>Chapter</td>
<td>Page</td>
</tr>
<tr>
<td>---------</td>
<td>------</td>
</tr>
<tr>
<td>4.2 Upward Gaze Implementation</td>
<td>27</td>
</tr>
<tr>
<td>4.3 Experimental Setup</td>
<td>29</td>
</tr>
<tr>
<td>4.4 Experimental Procedure and Conditions</td>
<td>32</td>
</tr>
<tr>
<td>4.5 Questionnaire</td>
<td>35</td>
</tr>
<tr>
<td>4.6 Data Processing and Analysis</td>
<td>36</td>
</tr>
<tr>
<td>4.6.1 Data Analysis</td>
<td>39</td>
</tr>
<tr>
<td>4.6.2 Behavior Encoding to Determine Gaze Following</td>
<td>39</td>
</tr>
<tr>
<td>5 RESULTS AND DISCUSSION</td>
<td>43</td>
</tr>
<tr>
<td>5.1 Effectiveness of the Virtual Environment</td>
<td>43</td>
</tr>
<tr>
<td>5.2 Behavioral Contagion</td>
<td>47</td>
</tr>
<tr>
<td>5.3 Conclusion</td>
<td>52</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>54</td>
</tr>
</tbody>
</table>
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>Virtual crowd simulation parameters</td>
<td>15</td>
</tr>
<tr>
<td>3.2</td>
<td>Virtual reality platforms’ specifications</td>
<td>19</td>
</tr>
<tr>
<td>4.1</td>
<td>Experimental conditions’ details</td>
<td>36</td>
</tr>
<tr>
<td>5.1</td>
<td>Participants’ average scores to survey questions on a 7-point scale</td>
<td>44</td>
</tr>
<tr>
<td>5.2</td>
<td>Probability of looking up sensitivity to the threshold</td>
<td>49</td>
</tr>
<tr>
<td>5.3</td>
<td>Proportion of time looking up sensitivity to the threshold</td>
<td>50</td>
</tr>
<tr>
<td>5.4</td>
<td>Proposed contagion model [21] variables</td>
<td>51</td>
</tr>
</tbody>
</table>
**LIST OF FIGURES**

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1 Top and front views of the developed virtual environment with crowd as part of the thesis.</td>
<td>2</td>
</tr>
<tr>
<td>2.1 Different force types in both social force and anticipatory collision avoidance models: goal force, interactin and wall force.</td>
<td>8</td>
</tr>
<tr>
<td>3.1 Experimental arena and its replication as a virtual environment</td>
<td>12</td>
</tr>
<tr>
<td>3.2 Samples of male and female virtual characters that have been used as virtual crowds.</td>
<td>13</td>
</tr>
<tr>
<td>3.3 Schematic of the VR platform developed for Samsung Gear VR using a combination of a phone, motion capture system, router and a computer.</td>
<td>18</td>
</tr>
<tr>
<td>3.4 Minimum processed frames per second during a simulation as a function of time step.</td>
<td>21</td>
</tr>
<tr>
<td>3.5 Minimum processed frames per second during a simulation as a function of interaction raduis.</td>
<td>22</td>
</tr>
<tr>
<td>3.6 Minimum processed frames per second during a simulation as a function of intervals between data points.</td>
<td>23</td>
</tr>
<tr>
<td>4.1 A sample character with selected bone joints on its body: spine 1 &amp; 2 and head bone.</td>
<td>28</td>
</tr>
<tr>
<td>4.2 Schematic of the experimental setup showing the silhouette of an actual participant.</td>
<td>31</td>
</tr>
<tr>
<td>4.3 Top view of participant’s initial position and colored cylinders placement in the virtual environment.</td>
<td>33</td>
</tr>
<tr>
<td>4.4 Pitch, roll, yaw and position trajectories of the participant from a sample trial with 5 stimulus characters.</td>
<td>37</td>
</tr>
<tr>
<td>Figure</td>
<td>Page</td>
</tr>
<tr>
<td>--------</td>
<td>------</td>
</tr>
<tr>
<td>4.5</td>
<td>38</td>
</tr>
<tr>
<td>4.6</td>
<td>41</td>
</tr>
<tr>
<td>4.7</td>
<td>42</td>
</tr>
<tr>
<td>5.1</td>
<td>45</td>
</tr>
<tr>
<td>5.2</td>
<td>46</td>
</tr>
<tr>
<td>5.3</td>
<td>47</td>
</tr>
<tr>
<td>5.4</td>
<td>48</td>
</tr>
<tr>
<td>5.5</td>
<td>49</td>
</tr>
</tbody>
</table>

4.5 Pitch, roll, yaw and position trajectories of the participant from a sample trial with 15 stimulus characters.

4.6 A snapshot of the opened video in BORIS software and table of key presses in right side of the window.

4.7 Accuracy curve across different thresholds for head pitch rotation.

5.1 Trajectories of participants and nearby virtual characters across all trials.

5.2 Participants’ moved distance as a function of their response to Question 10.

5.3 The probability of participants who looked up depended on the number of virtual characters who looked up.

5.4 Average of participants’ rate to Question 9.

5.5 Proportion of time that participants looked up depended on the size of the stimulus group, with more time spent as the number of virtual characters who looked up rose.
CHAPTER 1
INTRODUCTION

Humans are social species who have been living in groups for thousands of years. With growing populations humans experience crowd environments in their daily life, especially in urban societies. Living in groups results in direct and indirect interactions between individuals, leading to social phenomena commonly termed as collective behaviors. Because of such interactions, collective behaviors are hard to explain in terms of individual actions.

Social influence is a common phenomena which is defined as alteration in beliefs, thoughts or behaviors of an individual due to interactions with others [1]. Examples of social influence is collision avoidance, where individuals steer their movement to avoid hitting each other, and behavioral contagion, where individuals tend to follow the behavior of those within proximity.

The communication of information between individuals in human crowds plays a fundamental role in how a group of humans responds to sudden events [2]. A better understanding of how information is communicated within groups can help us determine the factors that play a role in crowd navigation and its response to rumors and sudden events. In this context behavioral contagion provides a visual indicator of how information flows within human crowds.

Behavioral contagion might end up with dangerous situations if it happens in very crowded environments (e.g. concerts, stadiums and festivals) and changes a crowd’s flow spontaneously. Considering different aspects of this phenomenon, understanding the nature and mechanism of it will help researchers in various areas, including social sciences, safety and evacuation scenarios, traffic management strategies, urban design and civil engineering.
Due to the impracticality and safety issues associated with conducting experiments with human crowds, the majority of research in crowd behavior has focused on the development of pedestrian motion models that are able to reproduce crowd motion qualitatively. In this context, virtual reality provides a viable alternative to perform laboratory experiments interfacing virtual and real humans to understand human response to different crowd situations.

In this thesis, we specifically aim to develop a realistic virtual environment consisting of a crowd of humanoid characters and use it to study the effect of behavioral contagion commonly present in real crowds. To do so, first an interactive and immersive virtual environment has been developed. This is accomplished by developing a pedestrian motion model using a combination of existing crowd models (Fig. 1.1).

Finally, a human behavior study was designed and conducted to evaluate the virtual reality platform. In this behavior study, we tested the hypothesis that the probability of the subject following an upward-looking gaze of the virtual characters is dependent on the
size of the stimulus virtual crowd. Participants responses to a questionnaire focused on their perceptions of the realism of the virtual environment were used to assess the performance. Experimental data of human head movement were processed and statistically analyzed to test the hypothesis that behavioral contagion was triggered in virtual environments. Results were compared to similar studies in the real world.

1.1 Background

Virtual reality (VR) has the potential to advance our understanding of human behavior in response to carefully controlled environments without the variability posed in natural settings [3]. Recent developments in computer graphics and animations have demonstrated the potential to elicit emotional contagions in participants when they were shown virtual characters exhibiting a variety of emotions [4] and to passive and active crowds like mobs [5].

Virtual reality as a fast-growing technology has a potential to ease experimental design and achievements in many ways. In virtual reality platform we can have a flexible, controllable and safe environment; we just need one participant instead of a large crowd; and it can be accurately controlled and repeated over several participants.

Virtual reality has been used in human behavior studies with different evacuation scenarios as well. Sharma and his team used VR to simulate aircraft emergency evacuation. In their simulation, users were able to navigate as virtual agents in the virtual environment and also guide the agents [6]. Kinateder and his collaborators tried to assess social influence effect in choosing evacuation routes in a virtual reality fire tunnel [7]. In another study, Moussaid and his team used VR immersive environment to investigate crowd behavior in a high-stress evacuation situation. They used 36 participants to model 36 agents who were
trying to evacuate a building and showed that their results are similar to typical patterns in real crowds [8].

1.2 Thesis Contribution

The purpose of this study is to investigate the effectiveness of virtual environments in triggering behavioral contagion, and the thesis uses a combination of complex system modeling, behavioral experiments and virtual reality technology to study social influence in virtual crowds. The contributions of this thesis are:

- A novel virtual environment that combines established models of pedestrian dynamics to simulate a realistic virtual crowd. In particular, we combine two collision avoidance strategies to ensure that virtual characters anticipate collisions from afar but are also able to form dense human crowds without touching when needed. The developed virtual environment is evaluated in terms of real-time performance and can simulate up to 60 virtual characters with high-resolution graphics. The virtual environment is generic and can be used to reproduce a variety of crowded situations, including the one used in the experimental study here.

- A hypothesis-based study to evaluate the existence of social influence in virtual reality. Specifically, we utilize the virtual environment to test the hypothesis that the probability of the subject following an upward looking gaze of the virtual characters is dependent on the size of the stimulus virtual crowd. Following the experimental procedure from an earlier study in social psychology, we elicit curiosity in participants by having a subset of a virtual crowd look in a particular direction. Post-experiment surveys are used to evaluate the believability of the virtual environment; postural re-
sponses are used to assess the extent of behavioral contagion induced in the virtual environment.

Material from this thesis has been used in the following peer-reviewed publication:

CHAPTER 2
CROWD BEHAVIOR MODELS

2.1 Pedestrian Dynamics

Pedestrian crowd is a complex system consisting of individuals interacting through multiple senses. By increasing researchers’ interest in studying human crowds, different crowd models have been developed. Mathematical modeling of pedestrian dynamics can be classified in terms of macroscopic and microscopic approaches.

In macroscopic approaches, a crowd is considered as a whole, behaving like a fluid. In this approach, behavioral factors are largely reduced and a major goal is to estimate specifications of the whole crowd, including velocity and density. In microscopic approaches, each individual is separately considered and given a degree of intelligence so they are more realistic. Microscopic approaches consider local interactions between individuals and also interaction between individuals and solid objects, which makes them effective methods in behavioral studies.

There are different particle-based crowd models with microscopic approach that have been used to model dynamics of a pedestrian crowd, including social force [9], anticipatory collision avoidance [10], optimal reciprocal collision avoidance [11], and centrifugal force models [12]. We chose and implemented two particle-based models, social force and anticipatory collision avoidance, in our study. Both implemented models will be described in detail in the following sections.
2.2 Social Force Model

Social force model [9] is a well-known force-based model which represents particles dynamic of motions in form of Equation 2.1. Interaction force between particles in this model is an exponential function of relative distance between them and increases as the particles get close to each other. In social force model, a sum of three different forces and noise($\eta$) is applied to each particle: goal force ($f_g$), interaction forces with all other particles ($f_{ij}$), and force with the walls or obstacles ($f_{iw}$) as illustrated in Figure 2.1.

$$m_i\ddot{r}_i = f_g^i + \sum_{j=1,j\neq i}^{N} f_{ij} + \sum_{W} f_{iw} + \eta$$

(2.1)

where the goal force is

$$f_g = \frac{s_i e_i(t) - v_i}{\mu_i}$$

(2.2)

Goal force represents the tendency to move towards a desired direction $e_i(t)$ with relaxation time $\mu_i$ in seconds [13], current velocity $v_i$ in m/s, and desired speed $s_i$ in m/s. By changing the desired direction, particles can move around and in different directions. The exponentially rising particle-particle interaction force is

$$f_{ij} = A_i e^{\frac{(r_{ij} - d_{ij})}{B_i}} n_{ij}$$

(2.3)

where $d_{ij}$ is two particles distance, $r_{ij}$ is summation of two particles’ radius $r_{ij} = R_i + R_j$, $A_i$ is interaction strength constant, $B_i$ is interaction range constant, and $n_{ij}$ is the unit vector along the relative displacement between the two particles. $f_{iw}$ denotes the interaction wall force which has the same form as the particle-particle interaction force but with static objects in the environment such as walls. In this case, walls and obstacles are considered as
Figure 2.1: Different force types in both social force and anticipatory collision avoidance models: goal force, interaction and wall force.

stationary particles in the environment that do not feel force from the dynamic particles but apply the same form of interaction force to them.

In this model, between each pair of characters there is always an interaction force, and collision-free movements are guaranteed, but continuously avoiding each other is not a realistic movement between pedestrians in some cases like pedestrians who are walking near and parallel to each other without any repulsive force between them.

2.3 **Anticipatory Collision Avoidance Model and Variants**

Anticipatory collision avoidance model follows the same equation of motion (2.1) and terminology as social force model but here the interaction force between agents is a function
of time to collision [10]. In this model, for each pair of particles, considering their relative
distance and velocity, time to collision, $\tau$, is calculated and according to it interaction force
is estimated. The time to collision is

$$\tau = \frac{b - \sqrt{d}}{a} \quad (2.4)$$

where $a = \|v_{ij}\|^2$, $b = -(x_{ij} \cdot v_{ij})^2$, $c = \|x_{ij}\|^2 - (r_{ij})^2$, $d = b^2 - ac$, $x_{ij}$ and $v_{ij}$ are relative
displacement and velocity. If there is a real time to collision between two agents, then the
repulsive interaction force between them is

$$f_{ij}^{ac} = -\left[\frac{k e^{-\frac{\tau}{\tau_0}}}{\|v_{ij}\|^2 \tau^2} \left(\frac{2}{\tau} + \frac{1}{\tau_0}\right)\right] \left[ v_{ij} - \frac{\|v_{ij}\|^2 x_{ij} - (x_{ij} \cdot v_{ij}) v_{ij}}{\sqrt{(x_{ij} \cdot v_{ij})^2 - \|v_{ij}\|^2 (\|x_{ij}\|^2 - (r_{ij})^2)}} \right] \quad (2.5)$$

$k$ is a scaling constant, $\tau_0$ is the largest value of time to collision at which collision must
begin to be avoided, and $r_{ij}$ as before is the sum of agent radii. In case that there is not a
real value for time to collision, it means that there is not a collision ahead and interaction
force is zero. This model promises collision-free movements for small enough time steps and
lets particles walk or stay near each other without any repulsive force when they are not
going to collide.

Anticipatory collision avoidance model assumes that each agent’s position and velocity
values are sensed exactly, while it is not a realistic assumption if it is implemented to be
used in devices with sensors. Possible sensor or calculation errors might add uncertainties
to the model to some extent.

These uncertainties have been addressed in the extended model named adversarial model.
Specifically, it was assumed that position and velocity errors are bounded to known values
$\delta$ and $\epsilon$ respectively [14]. Adversarial model is a version of anticipatory collision avoidance.
model that conservatively assumes these possible errors are towards direct collisions and changes the relative velocity between agents considering the error bound as follows:

\[
\tilde{v}_{ij} = v_{ij} - \epsilon \frac{x_{ij}}{\|x_{ij}\|} \tag{2.6}
\]

This model assumes that the sensed velocity differs from the actual velocity and corrects it based on the error bound value. Therefore, the time to collision between agents is calculated using this changed relative velocity. The interaction force is estimated by substituting time to collision and changed relative velocity, \(\tilde{v}_{ij}\), in the same interaction force formula that is used for the anticipatory collision avoidance model.
CHAPTER 3
VIRTUAL REALITY ENVIRONMENT

The purpose of this study is to evaluate virtual reality technology as a viable platform for studying crowd behavior in a virtual environment. In this context, virtual characters who walk through an indoor arena using an adapted motion model play the role of the stimulus group of actors as they stop and look up, and an individual participant’s position and orientation are observed as he or she interacts with them.

3.1 Virtual Environment Design

The virtual environment was authored in a 3D graphics software (Unity, version 2017.1.0f3). Static features within the environment such as walls, desk arrangements, and whiteboards were created to replicate the architecture and geometry of an actual environment, namely the robotics lab within the engineering building of Northern Illinois University (Fig. 3.1).

Dynamic features within the environment included 60 humanoid characters consisting of 32 males and 28 females. These were generated from 16 unique male and 15 unique female characters obtained from the Unity Asset store (Modern People and Modern People 2 packages, independent developer) by changing their dress colors. Samples of virtual characters are shown in Figure 3.2.

The humanoid characters were rigged and skinned and each had a set of 55 body joints allowing them to exhibit human-like movements such as walking and turning. All the char-
Figure 3.1: Experimental arena (a) where the experiments were conducted. Participant’s perspective of the virtual environment (b) that was authored to match the real environment by adding distinct features such as desks, whiteboards, cabinets and display monitors at similar locations. Top view of the virtual environment (c) with sample trajectories marking the walking route of virtual characters.
acters were in FBX format and their animation type must have been selected as Humanoid in import settings so that they could present human-like movements.

Figure 3.2: Samples of male and female virtual characters that have been used as virtual crowds.

Animations of walking and standing idly were obtained separately from the Unity Asset store and assigned to each character. Head movement for modifying gaze direction within experimental conditions was authored and controlled using an additional plugin [15] (Final IK, rootmotion, Tartu, Estonia).
3.2 Virtual Crowd Simulation

Our initial simulations showed that while social force model guaranteed collision-free simulations, the anticipatory collision avoidance model displayed a more natural avoidance at larger distances. Accordingly, the interaction between the characters was simulated as a combination of both models to achieve the most visually realistic motions.

In particular, when humanoid characters were not in direct contact, they avoided collisions on the basis of time to collision [10] while accounting for uncertainty in position and velocity estimates among individuals [14]. If the characters came in direct contact, they experienced a distance-based force [9] to keep them apart. Each character is approximated by a finite-sized circle and the interaction force between characters is calculated if they are within 4 meters distance from each other to improve the simulation’s efficiency. Then, the combined crowd model is

\[ f_{ij} = \begin{cases} f_{ac}^{ij}, & \text{if } \|x_{ij}\| > r_{ij} \\ f_{sf}^{ij}, & \text{if } \|x_{ij}\| \leq r_{ij} \end{cases} \]

where \( r_{ij} \) is sum of their radii and is equal to 0.8 m. The details of the models have been explained in Sections 2.2 and 2.3.

Interaction force between characters and walls or stationary objects in the virtual environment can be obtained using either social force or anticipatory collision avoidance model with the same approach since walls are considered as stationary characters without humanoid shape. We used social force model to calculate all interaction forces between characters and walls since it is cheaper in terms of computational costs. These objects do not feel the opposing force from the characters themselves.
Uncertainty in the environment and movement is captured by the noise term sampled from a uniform distribution $\eta \in U(5,10)$ Newtons at approximately 10% of the maximum force observed in the simulation.

It must be considered that Unity has a default collision detection option for each game object in the scene that must be disabled; otherwise, it might intervene with the models. Table 3.1 lists the parameters used to simulate the virtual crowd.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$m_i$</td>
<td>mass of each agent</td>
<td>60 kg</td>
</tr>
<tr>
<td>$s_i$</td>
<td>desired speed of each agent</td>
<td>0.7 m/s</td>
</tr>
<tr>
<td>$\tau_i$</td>
<td>relaxation time for each agent</td>
<td>0.25 s</td>
</tr>
<tr>
<td>$A_i$</td>
<td>interaction strength for each agent</td>
<td>1000 N</td>
</tr>
<tr>
<td>$B_i$</td>
<td>interaction range for each agent</td>
<td>0.08 m</td>
</tr>
<tr>
<td>$\tau_0$</td>
<td>max. time-to-collision to enable avoidance</td>
<td>5 s</td>
</tr>
<tr>
<td>$k$</td>
<td>scaling constant for each agent</td>
<td>1.5</td>
</tr>
<tr>
<td>$\epsilon$</td>
<td>uncertainty factor in velocity estimate</td>
<td>0.2</td>
</tr>
</tbody>
</table>

Character animations were controlled through the crowd motion simulation by directly updating animation variables such as stepping speed, transition to a stop state, and a trigger to alter the gaze direction. This ensured that a character moving fast took faster steps and did not exhibit unnatural leg movements when it was supposed to stop. Triggers to alter the gaze direction were based on character position relative to the participant’s avatar.

### 3.3 Virtual Crowd Trajectories Within the Indoor Lab

To create the appearance of a crowd flow, the virtual characters were made to walk through twelve real-time generated waypoints within the virtual environment. The waypoints were selected on ellipses so that the characters entered the environment through one of the exits and left through the other. These exits corresponded to actual exits of the lab.
The first goal point for each virtual character, $i$, is $[r_{ix} \cos \theta_{i1}, r_{iz} \sin \theta_{i1}]$ while $r_{ix}$ and $r_{iz}$ are randomly picked from a specified range and $\theta_{i1} = \arctan \left( \frac{x_0}{z_0} \right)$. The characters walk toward their first assigned goal point and once they reach within 1 m of that goal point a new one will be generated as follows:

$$r_x \in \mathbb{U}(c_1, c_2), r_z \in \mathbb{U}(c_3, c_4)$$  \hspace{1cm} (3.2)

$$\theta_{k+1} = \begin{cases} 
\theta_k + \Delta \theta, & \text{if clock-wise walking} \\
\theta_k - \Delta \theta, & \text{if counter-clock-wise walking}
\end{cases}
$$

where $\Delta \theta = 2\pi/N_{\text{waypoints}}$ and $c_1, c_2, c_3, c_4$ are constants that specify the available range that characters can walk in. To give a natural appearance of a moving crowd, approximately 30% of the characters were set to walk in the opposite direction. A few samples of virtual characters’ trajectories are shown in Fig. 3.1 part (c).

### 3.4 Virtual Reality Platform

Virtual reality as a fast-growing technology is being used in diverse range of applications and has been developed in different types of platforms. Head-mounted display (HMD), a desktop monitor and virtual cave are different forms of virtual reality platforms that might be selected for an application.

We needed a platform that provides a large walkable area to model a virtual crowd and study spatial interactions. In addition, we preferred a wireless or tether-free system to ease participant movements and simultaneously increase the walking range.

Head-mounted display systems are widely known and popular due to the immersiveness that they offer. Considering the details of our study, we needed to choose a platform that
lets participant walk and also provide a very immersive experience, so we decided to use a head-mounted display platform in our study. In this section we will explain which virtual reality platform we chose to use in our study and the consequent technical challenges of the platform.

### 3.4.1 Samsung Gear VR

Most famous virtual reality platforms like Oculus Rift and HTC Vive can only provide a limited area for user to walk in. In addition, in these platforms the headset is attached to a computer by a long wire. We preferred a larger tether-free area for our study, which was not available. Therefore we decided to use Samsung Gear VR and its compatible mobile phone to develop a tether-free platform.

Samsung Gear VR is a head-mounted display designed for Samsung mobile phones and plays the virtual reality products (e.g. movies, games) which are available on the phone. We made an executable format of our virtual simulation to be installed and played on the phone. As this platform worked with an integrated phone, the user could freely walk when wearing it. The important point was that if the user turned his head, his view got updated using the phone gyroscope’s data, but if he walked, the virtual environment did not get updated accordingly. For example, if the user went toward an object in the virtual environment, he did not get closer to it as he expected. This was happening because user’s position data did not get updated in the virtual environment.

In order to fix this issue we used a motion capture system, OptiTrack Trio camera, to track users movements. The motion capture system was able to stream tracked data to other devices. A Python library [16] was used to receive streamed data from the camera and write it in a text file on a computer. This text file was transferred to the phone’s SD card
through File Transfer Protocol (FTP) so that the virtual simulation reads user’s position data from it. The whole streaming, receiving, writing and reading processes between the three devices were done in real time, which ended up with a huge delay (approximately 2 seconds) on position updates in the virtual environment. We noticed that this approach was not practical and switched our method. The schematic of the setup is shown in Fig. 3.3.

Figure 3.3: Schematic of the VR platform developed for Samsung Gear VR using a combination of a phone, motion capture system, router and a computer. The sequence of tasks in the loop is shown in order.

### 3.4.2 HTC Vive

HTC Vive is a virtual reality platform which includes a head-mounted display, two trackers and two controllers. The virtual simulation must be played on a computer while the headset is connected to the same computer with a long wire. In this setup, the simulation is executed on the computer and presented for the user on the headset. The user will be able to walk within the specified area that is covered by the trackers range.
This platform became available commercially during the time we were developing the Samsung Gear VR setup. It became a well-known product soon and we switched to that as our developed setup was not practical in our study. HTC Vive and Samsung Gear VR specifications are presented in Table 3.2.

<table>
<thead>
<tr>
<th></th>
<th>HTC Vive</th>
<th>Samsung Gear VR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Resolution</td>
<td>1080 × 1200 (per eye)</td>
<td>2560 × 1440</td>
</tr>
<tr>
<td>Refresh rate</td>
<td>90 Hz</td>
<td>60 Hz</td>
</tr>
<tr>
<td>Processor speed</td>
<td>2.9 GHz (Separate PC)</td>
<td>2 × 2.15 + 2 × 1.59 GHz (Phone)</td>
</tr>
<tr>
<td>Price</td>
<td>$799 + $1000 (System requirements)</td>
<td>$100 + $700 (Phone cost)</td>
</tr>
<tr>
<td>Delay</td>
<td>22 ms</td>
<td>2 s</td>
</tr>
<tr>
<td>Mobility</td>
<td>Tethered</td>
<td>Tether-free</td>
</tr>
<tr>
<td>Weight</td>
<td>1.2 lb (without cables)</td>
<td>1 lb (with phone)</td>
</tr>
</tbody>
</table>

To make the walking more convenient for the user, we decided to replace the computer and the long wire with a backpack computer and a short wire respectively. This backpack computer was a compact computer without a desktop that the user could wear and walk without tethering concerns.

HTC Vive’s tracking precision is high and its latency is very low (22 ms), though the reported positions and orientations have inaccuracies to some extent [17]. Specifically, Vive’s internal coordinate system is tilted with respect to the physical reference ground plane, which greatly affects the height measurements. In addition, due to the tilted reference plane, roll and pitch measurements will have inaccuracies as well. However, the authors in [17] showed that inaccuracies in the measurements except in the vertical dimension are in acceptable range as long as the tracking is not lost. System calibration prior to usage is suggested to minimize the errors in measurements.
3.5 Real-Time Performance

In virtual reality simulations the level of presence and immersion highly depends on display refresh rate; a higher frame rate results in more immersive experience and less motion sickness. The HTC Vive headset is capable of display at 90 Hz but the processing time could affect this rate. We aimed for maintaining the simulation frame rate at minimum 60 Hz and managed the whole process accordingly.

We did some real-time performance analysis to investigate frame rate dependency on different simulation factors. In our case, the simulation frame rate was significantly affected by crowd model simulation time step, interaction range, and data collection sampling rate, which must be chosen carefully. Therefore, we measured and plotted the simulation’s least performance in terms of frames per second for different values of the aforementioned variables.

Fig. 3.4 shows the minimum frame per second during a simulation as a function of time step. It should be noted that the frame rate values are the least values that happened instantly in an entire simulation and the average values are higher. As mentioned before, performance of anticipatory collision avoidance model and its variants depends on time step value: the smaller values guarantee the better performance. Although it was a challenging and computationally expensive decision, we used a very small time step, 0.02 seconds, in the crowd simulation.

The small time step that we used resulted in a heavy load on the platform, so needed to significantly reduce other processing times. To do so, we did script optimizations with replacing time-consuming commands and algorithms with more efficient alternatives, avoiding multiple memory allocations, simplifying mathematical calculations and applying recommended settings for reducing rendering and graphical time process.
Figure 3.4: Minimum processed frames per second during a simulation as a function of time step.
The aforementioned changes were still not enough to achieve consistent 60 Hz fresh rate for the head-mounted display. In the next round of optimization, we noticed that our data collection process was the bottle neck. We were collecting required data for each frame and wrote it on the text file in real time. We defined a variable called interaction range to collect our required data within a specified range from the participant and also changed the intervals between data collection. The effect of the mentioned factors on the performance are shown in Figures 3.5 and 3.6 respectively.

![Figure 3.5: Minimum processed frames per second during a simulation as a function of interaction radius.](image-url)
Figure 3.6: Minimum processed frames per second during a simulation as a function of intervals between data points.
Finally, the data collection was limited with 4 meters interaction range and performed every 10 frames. In addition, the collected data wrote in a text file at the end of trial instead of in real time.

The implemented model with the small time step resulted in collision-free movements but not smooth enough, especially on turns. Specifically, characters’ change of direction were sudden and jerky when they avoided each other. Considering the actual human movements, we found out that rotation rate must be a function of characters’ velocity. Therefore characters’ rate of turning from their current direction towards their target direction was set proportional to their speed and consistent smooth movements were achieved.
CHAPTER 4

HUMAN-SUBJECTS STUDY ON BEHAVIORAL CONTAGION

The experimental procedure followed in this study was approved by the Institutional Review Board at the Office of Research Compliance, Integrity and Safety under the protocol # HS17-0287. Seventy subjects (51 male and 19 female, age of all participants = 23.4 ± 4.24 years) were recruited for the study through flyers posted in the Northern Illinois University buildings. Subjects were above 18 years of age who had no history of seizures or motion sickness from computer displays. Subjects were asked to review and sign a consent form with an additional consent for allowing to film an overhead view. Experiments were conducted between November 2017 and April 2018.

4.1 Behavioral Contagion Experimental Study

Behavioral contagion, as a form of social influence, plays a fundamental role in how animal groups respond to new information [18], [19]. Considering the effect of this phenomenon on human decision making, it is of great importance to know how behavioral contagion is triggered. In some cases, behavioral contagion may significantly change the dynamics of a crowd and end up with dangerous and uncontrollable situations. Examples include disasters that take place in crowded environments like concert halls, stadiums, and religious traditions.

Modeling and studying behavioral contagion in different contexts helps to understand triggering factors and develop control strategies. In addition, behavioral contagion changes
the aspects of a crowd that can be measured and quantified in engineering contexts, e.g. crowds’ flow, density, and velocity.

Behavioral contagion literature shows that it has been studied in multiple outdoor experiments. The common difficulty between these studies is that outdoor experiments generally have inaccuracy and uncertainties about people’s behaviors. In addition, they usually have to control and monitor a crowded environment during the experiment, and data collection gets challenging accordingly. In one of the earliest studies about behavioral contagion, a study was performed on a busy street in New York City to study the size effects of the crowd on it [20]. In that study, a set of actors stopped and looked up at a specific window of a building. Results from the study showed that the number of pedestrians mimicking the gaze increased with the size of the stimulus crowd of actors and plateaued beyond a certain number. Gallup and colleagues [21] tried to reproduce these results; they could not get the same quorum-like behavior, but they showed the same stimulus group size effect on passers’ response. They also showed that the social response itself was context dependent even between two different day-to-day scenarios: a train station and a shopping street. In another study related to gaze following, they used the same scenario to study the effect of directional flow on visual information transfer as well [22]. They observed that people had more propensity to look up when they were behind the stimulus rather than in front of them, in other words, when they were not watched.

There are other studies that showed the same effect that Milgram et al. [20] obtained in different contexts. Latane and Darley showed that bystander effect increases when there are more passive individuals in a smoke-filled room. Subjects were less likely to consider the situation dangerous in presence of passive individuals [23]. In another study, Leon Mann conducted an experiment to test the stimulus size effect on queue-joining behavior of commuters in a bus stop. They asked actors to queue in a spot where it was not customary for queuing and were able to achieve a considerable queue joining level after a specific stimulus
number [24]. Freedman et al. studied the effect of density and number of stimuli group on applauding after a movie [25]. Richard Mann et al. have also studied the spread of applause after an academic presentation and showed that after some persons started to clap, others joined them [26]. They noticed that both applause starting and cessation were contagious, but cessation is a bit controlled by the audience’s unwillingness to clap too much. In contrast to other studies in this area, they observed in their experiment the spatial vicinity was not important, probably due to the acoustic nature of clapping.

Considering these studies, we decided to investigate the probability of the subject following the virtual characters’ upward gaze as a function of the stimulus size. We elicited curiosity in participants by having a subset of a virtual crowd look in a particular direction. Behavioral contagion was quantified by recording the postural response of the participant as they explored the virtual environment.

### 4.2 Upward Gaze Implementation

As part of the experiment scenario, some of the characters were supposed to stop and look up in a predefined situation. Humanoid characters have a system of bones and joints that are controlled by either animations or scripts (Figure 4.1). Common human movements (e.g. walking, running, jumping, etc.) are usually presented using animations, and rare movements are applied on top of the existing animations as minor changes. In our case, we required some characters to:

1. Stop walking, which was accomplished by transitioning between two animations (walking to staying idle)

2. Look up, which was accomplished by turning the character’s neck bone upward while the idle pose animation is active simultaneously
3. Keep looking up for a specified time

4. Look forward

5. Resume walking

![Figure 4.1: A sample character with selected bone joints on its body: spine 1 & 2 and head bone.](image)

Transition between animations is controlled by character's speed in the whole simulation but the challenging part is looking up. To do so we used a Unity plugin, Final IK [15], which is a powerful tool to apply minor changes to the existing animations. This plugin makes scripting for animation changes easier and smoother. There are different scripts with different applications in this package; we used AimIK script based on our application.
AimIk takes a selected set of bones based on the desired movement to rotate the specified joint bone towards a target. We created an invisible looking target which was 12 m above the head of the participant. The selected bones in our case were Head and Spine2 joints of humanoid characters due to the nature of the looking up mechanism (Figure 4.1). Each selected bone has a weight with zero default value; by changing the weights the specified joint (head transform) will rotate towards the target transform (looking target) with the following relations:

\[
\begin{align*}
H_{k+1} &= H_k + W_1 dt, \quad t < t_s \text{ and } H_k < H_{\text{max}} \\
S_{k+1} &= S_k + W_2 dt, \quad t < t_s \text{ and } H_k < H_{\text{max}} \\
H_{k+1} &= H_k - W_3 dt, \quad t > t_s \text{ and } H_k > H_{\text{min}} \\
S_{k+1} &= S_k - W_4 dt, \quad t > t_s \text{ and } H_k > H_{\text{min}}
\end{align*}
\]

where \(H\) and \(S\) are Head and Spine2 bone weight values, \(t_s\) is the time duration that stimuli are triggered, and \(H_{\text{max}}\) and \(H_{\text{min}}\) are maximum and minimum head bone weight thresholds. \(W_1, W_2, W_3\) and \(W_4\) are constant rotation rates which were assumed to be 0.08, 0.15, 0.5 and 0.8 in our simulation respectively.

This procedure is applied only to the selected characters as stimuli. These characters are selected in real time according to their proximity to the participant. At a random instant after a predefined time the nearest virtual character to the participant will be selected as stimulus and within a specified interval the next nearest virtual character to her or him will be selected till the specified stimuli number is achieved.

### 4.3 Experimental Setup

The experimental setup which is presented in Fig. 4.2 consisted of:
1. a head-mounted display (HTC Vive, with 1080 × 1200 pixel displays sampled at 90 Hz) integrated with motion trackers (Lighthouse laser emitters) to present a walkable virtual environment

2. a backpack computer (MSI VR one backpack PC, 2.90 GHz Processor, 16 GB memory, 3.3 kg) for running the virtual environment simulation

3. a pair of noise-cancelling headphones (Bose QuietComfort acoustic noise cancelling headphones)

4. a laptop (Laptop I, Dell Latitude) to initiate the environment simulation and record the participant’s perspective of the virtual environment using a remote desktop connection

5. a laptop (Laptop II, Dell Vostro) for recording an overhead view of the participant movements in the real world

6. and a web camera (Logitech C920 Pro Webcam) to record overhead view of the participant

The experiments were conducted in the robotics lab located on campus. The lab is 18 m long × 9 m wide × 4 m high. Within the lab, a 4.8 × 4 m area was allocated for experiments. Within this region, the motion trackers were placed diagonally 4.8 m apart as per manufacturer recommendations to track participant movement within a 3 m × 4 m area. The trackers were fixed in one place throughout the study and recorded the position and orientation of the HMD at 120 Hz.

In order to address the reported inaccuracies in HTC Vive position and orientation measurements (Section 3.4.2), the following considerations had been made:

- The height measurements were not collected due to the considerable errors.
Figure 4.2: Schematic of the experimental setup showing the silhouette of an actual participant. The backpack computer allows tether-free movement in the space between the motion trackers. The backpack computer is controlled with Laptop I; Laptop II is used to record an overhead view of the experiment. A coordinate frame attached to the HMD denotes how the roll, pitch, and yaw were measured.
• The HMD and motion-capture system was re-calibrated before the first trial each day that experiments were conducted.

• Participants were informed to stay within the experimental arena so that tracking was never lost.

• The trials were initiated from an approximate same location.

4.4 Experimental Procedure and Conditions

At the beginning of each trial, the experimenter explained the features of the virtual environment and platform to participants prior to asking them to wear the headset and backpack computer. These included detailing the equipment, the tracking cameras, the range of tracking and informing them that they will receive further instructions from the headset once the experiment is started. A sign was placed near the arena to ensure no one interrupted the experiment while it was being conducted. The participants were allowed to adjust the HMD, focus the display, and the headphones until they were comfortable.

Each participant was requested to put on the headset and the backpack while standing at the same approximate location within the arena. Once the participant was ready, the experimenter left the arena to take position behind a workbench and observe the participant. Although never required, the experimenter was ready to intervene if the participant were to collide with an object. At this time, the experimenter started playing the virtual environment for the participant and initiated recording of the participant perspective as well as an an overhead view of the experimental arena for subsequent analysis.

The participant was presented with a scene resembling the experimental arena in terms of the geometry and arrangement. The participants could not see their avatar’s hands. In
addition, two cylinders (colored green and blue) were placed 2.6 m apart, with the participant’s starting position 1.7 m away from each cylinder so that the three locations formed a triangle (Fig. 4.3).

![Figure 4.3: Top view of participant’s initial position and colored cylinders placement in the virtual environment. The black circle indicates the participant’s initial position.]

There was a habituation phase in the beginning of the trial consisting of the participant being instructed to walk between the two cylinders, starting with the blue cylinder. The first instruction was given after 15 seconds and once the participant reached within 0.5 m of the blue cylinder and 25 seconds had passed after the first instruction, a second instruction was given to walk towards the green cylinder. The instructions were recorded by a native speaker of English to ensure ease of understanding. No virtual characters entered the arena during this phase.

Virtual characters start walking into the arena after the second instruction, after approximately 40 seconds. The spatial movement of the characters was based on the interaction force defined in Section 3.2. In addition to avoiding collisions among themselves and with static objects in the environment including the two cylinders, the characters also avoided collisions with the participant avatar as they moved within the virtual environment. While the virtual characters were walking around the virtual environment, two more instructions
were repeated with the same time interval between instructions as soon as the participant reached within 0.5 meter of the cylinders.

After the third instruction, at a random time within the next 30 seconds, which was approximately 100 seconds from the beginning of the experimental trial, the stimulus was triggered. A subset of the virtual characters were selected based on their vicinity to the participant in real time and then stopped and shifted their gaze to look up slowly. Specifically, at a random time the virtual character nearest to the participant stopped and shifted his gaze to look up at a point 12 m high on the wall behind and above the participant’s head. There was nothing peculiar at this point other than the ceiling of the virtual room meeting the wall. Following this, the next nearest character also stopped and looked up after a delay of 0.5 seconds so that a cascade effect of looking up was observed. The delay of 0.5 seconds corresponded to delay due to observation and action [27]. This continued until the total number of characters selected to look up for that particular trial was achieved. The virtual characters maintained their upward-looking gaze for approximately 50 seconds. During this time other characters continued to walk while avoiding collisions with those who had stopped as well as the participant’s avatar.

In a between-subjects design, we varied the size of the stimulus group based on the number of virtual characters that looked up. Accordingly, the experimental conditions consisted of 1, 2, 3, 5, 10, and 15 characters looking up. The stimulus group for a participant was randomly selected so that we had approximately the same number of participants for each condition. The trial lasted approximately 4 minutes, at which time the participant saw a blank screen. Participants were then requested to remove the headset and the backpack computer and fill out a questionnaire, which will be explained in next section.

We recorded the following variables to quantify participant response to the virtual environment: position and head orientation of the participant plus the identity, position, and
head orientation of all the virtual agents within a 4 m distance of the participant throughout the trial.

4.5 Questionnaire

The questionnaire asked ten questions each ranked on a 7-point scale. The questions were designed to capture the degree of presence [28], motion sickness effect, participants’ evaluation about the environment and whether they observed any characters looking up. The questions are listed below.

1. How would you describe your past experience with virtual reality?

2. How natural did you find the movement of the characters?

3. How responsive did you find the virtual environment?

4. How well could you walk towards the cylindrical targets?

5. How crowded did the environment feel?

6. How comfortable did you feel?

7. The extent to which you felt as if you were moving when standing still?

8. The extent to which you felt the characters were reacting to your presence?

9. Did you notice any of the characters stop and look up?

10. How interested were you to explore the virtual environment or interact with characters?
4.6 Data Processing and Analysis

We recorded the two-dimensional position and three-dimensional head orientation (roll, pitch, and yaw) of the participant as available from inertial measurement sensors integrated into the HMD. We also recorded the position and gaze status (looking up or not) of all the virtual characters within a specified range from the participant in the virtual environment.

The aforementioned data were collected during the trial and written in a text file at the end of each trial. This procedure significantly affected processing time and we had to manage it in order to minimize the lag and maintain a refresh rate of at least 30 Hz for the participant. Therefore data were collected every 10 frames (10 Hz) and also the specified range for collecting virtual characters’ data considered 4 meters from the participant, which ensures the whole experimental arena coverage. In this setting, all the virtual characters’ positions and gaze status within the experimental arena were collected.

Seventy trials were done and data files were collected. Data of two subjects had to be removed from the analysis: one where the audio did not work properly and another where the audio instructions were misunderstood, leading to inactivity. The final number of subjects were 68 (49 male and 19 female). Different conditions and number of trials per condition are represented in Table 4.1.

<table>
<thead>
<tr>
<th>Condition</th>
<th>#of stimulus</th>
<th>#of trials</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>12</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>12</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>10</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>12</td>
</tr>
<tr>
<td>5</td>
<td>10</td>
<td>11</td>
</tr>
<tr>
<td>6</td>
<td>15</td>
<td>11</td>
</tr>
</tbody>
</table>
Prior to analysis, the data were detrended to remove the average position and orientation values across the full trial to compensate for any bias due to preferred head orientation and offsets [17] and re-referenced with respect to the first sample. Two sets of sample trial data are depicted in Figures 4.4 and 4.5.

Figure 4.4: Pitch, roll, yaw and position trajectories of the participant from a sample trial with 5 stimulus characters. The instant when the virtual characters looked up is marked by a vertical red line.
Figure 4.5: Pitch, roll, yaw and position trajectories of the participant from a sample trial with 15 stimulus characters. The instant when the virtual characters looked up is marked by a vertical red line.
4.6.1 Data Analysis

In order to quantify participants’ responses to virtual stimulus, it has been tried to answer specific questions about subjects’ reactions to virtual characters’ looking up. To do so, one minute before and after the stimulus is triggered, which is defined by the instant when the number of virtual characters looking up has reached the allocated value, were observed and the following questions were answered as explained:

1. Did they look up? The value of head pitch rotation is computed, and if the value was less than the threshold and the minimum value over the two minutes of observation was in the one minute after the stimulus, then yes.

2. How long did they look up? The total time the head pitch rotation was below the threshold was considered as their looking up duration.

In addition, participants’ answers to the questionnaire were encoded and used as a sort of validation when they were applied.

Considering the aforementioned assumptions, the collected data were statistically analyzed using MATLAB statistics toolbox. In particular, we used statistical analysis to investigate the effects of the independent variables (stimulus size) on the dependent variables (e.g. probability of looking up and proportion of time looking up).

4.6.2 Behavior Encoding to Determine Gaze Following

As a first step of data analysis, the looking up behavior had to be defined. A manual behavior encoding procedure based on head pitch orientation was used to find the threshold. Two sets of behavior encoding have been done to find a threshold for looking up behavior.
Specifically, we randomly selected 14 overhead videos of the participants in the real world and had them encoded twice by five independent observers. The observers were asked to watch these videos carefully and press a specific button on the keyboard as soon as the participant looked up.

The videos were assigned so that each video was encoded twice by two different observers. The encoding was performed in the BORIS software [29] that captures pre-assigned keystrokes marking the times when the observer recorded a look up by a participant (Fig. 4.6). Instances of look-ups that were encoded by two observers independently within a two-second delay (to account for attentional and keypress delays) were marked as a confirmed look up.

The confirmed look-up instances were then used as ground truth to estimate the threshold for head pitch orientation based on the following algorithm. Let the number of trials that have been encoded twice be $N_{tr}$. For a trial $tr = 1, \ldots, N_{tr}$, we need to find consistent look-ups between two encoding instances. Here a consistent look-up is defined as one that is within 0.25 seconds of each other (human reaction time). To find these, a sequence of time $(t = [0, 300])$ with a $\Delta t = 0.25$ seconds and two copies of one-dimensional encoding vectors of zeros whose size is the same as $t$ was created. For the first encoding vector ($ve_1$), each time the first encoder marked a look-up in an interval, the value of the encoding vector at that location is set to 1; for the second one, each time the second encoder marked a look-up the location in the encoding vector ($ve_2$) is set to 1. The elements that are equal to 1 in the element-wise product of these encoding vectors ($vc = ve_1 \cdot ve_1$) corresponds to confirmed look-ups.

To find the optimal threshold, the same approach was followed. Another one-dimensional vector of zeros with the same size as $t$ was created. Let a threshold for pitch direction be $T_p < 0$ and $\phi(t)$ represent instant pitch value. A look-up in the pitch data is counted if the
pitch is below the threshold $\phi(t) < T_p$ within a 0.25-second interval. Any time $\phi(t) < T_p$, the corresponding location in the pitch vector ($vp$) is set to 1.

Now we have two binary vectors across the time: confirmed look-ups vector as the ground truth ($vc$) and head pitch data vector ($vp$). We used binary classification to calculate accuracy for the given threshold. Accuracy is defined as

$$ACC = \frac{TPR + (1 - FPR)}{\text{Total confirmed look-ups}}$$

(4.1)

where $FPR$ and $TPR$ are false positive and true positive rates respectively. A false positive denoted that the head pitch orientation crossed the threshold within a two-second window, when no confirmed look-up was recorded by the observers, and a true positive denoted that a confirmed look-up was recorded at the time when pitch orientation crossed the threshold.

This algorithm has been repeated for threshold values $T_p \in [-30, -10]$ to find the threshold that maximized the true positive rate and minimized the false positive rate [30].
The threshold was found to be -17 degrees with 90.1% accuracy (Fig. 4.7). To check if our results were sensitive to any measurement errors inherent in the hardware [17], we conducted a sensitivity analysis of ±2 degrees about this threshold.

Figure 4.7: Accuracy curve across different thresholds for head pitch rotation.
CHAPTER 5
RESULTS AND DISCUSSION

5.1 Effectiveness of the Virtual Environment

In order to assess the effectiveness of the virtual environment, we analyzed participants’ answers to the questionnaires. Seven questions out of ten (Questions 2–4, 6–8, and 10) were focused on assessing the level of presence [28] based on their perception of how natural and responsive the environment felt and whether they experienced any discomforts due to visual delays. Participants’ average scores to each question of the survey are listed in Table 5.1.

Participants’ scores to the survey (on a 7-point scale) shows that they generally did not have prior experience with virtual reality. They evaluated both characters’ natural movements (3.8 ± 1.42) and their responsiveness (3.85 ± 1.9) as moderate, though most of them were not trying to get engaged with the virtual characters, which were interactive in terms of collision avoidance. In addition, participants were comfortable (5.4 ±1.5) and interested (5.9 ± 1.36) in the virtual environment and able to walk easily in the environment and had little indication of visual delays. These scores show that the participants were satisfied with the environment and crowd simulation as well, though possible improvements in crowd models and also computational power would make the simulation more satisfactory.

The participant and virtual characters trajectories in the experimental arena show that participants freely explored the virtual environment and also engaged with the virtual crowd. All trials’ trajectories are depicted in Fig. 5.1.
Table 5.1: Participants’ average scores to survey questions on a 7-point scale

<table>
<thead>
<tr>
<th>Question</th>
<th>Scale anchors</th>
<th>Mean±std</th>
</tr>
</thead>
<tbody>
<tr>
<td>How would you describe your past experience with virtual reality?</td>
<td>Very rare</td>
<td>Occasional</td>
</tr>
<tr>
<td>How natural did you find the movement of the characters?</td>
<td>Very artificial</td>
<td>Borderline</td>
</tr>
<tr>
<td>How responsive did you find the virtual environment?</td>
<td>Not responsive</td>
<td>Moderately responsive</td>
</tr>
<tr>
<td>How well could you walk towards the cylindrical targets?</td>
<td>With great difficulty</td>
<td>With some difficulty</td>
</tr>
<tr>
<td>How crowded did the environment feel?</td>
<td>Very crowded</td>
<td>Somewhat crowded</td>
</tr>
<tr>
<td>How comfortable did you feel?</td>
<td>Very uncomfortable</td>
<td>Somewhat comfortable</td>
</tr>
<tr>
<td>The extent to which you felt as if you were moving when standing still?</td>
<td>Not at all</td>
<td>Somewhat</td>
</tr>
<tr>
<td>The extent to which you felt the characters were reacting to your presence?</td>
<td>Not responsive</td>
<td>Moderately responsive</td>
</tr>
<tr>
<td>Did you notice any of the characters stop and look up?</td>
<td>Did not notice</td>
<td>Noticed a few</td>
</tr>
<tr>
<td>How interested were you to explore the virtual environment or interact with characters?</td>
<td>Not interested at all</td>
<td>Somewhat interested</td>
</tr>
</tbody>
</table>
Moreover, participants’ responses to Question 10 of the survey, which asked the degree to which they were interested in exploring the virtual environment, correlated linearly with the total distance walked during the experiment (linear regression test, $r = .411, p = .0005$).

Fig. 5.2 shows participants’ moved distance in the virtual environment as a function of their self-reported interest in the virtual environment.
Figure 5.2: Participants' moved distance as a function of their response to Question 10.
5.2 Behavioral Contagion

**Probability of looking up**: Figure 5.3 shows the probability of looking up by the participant varied with the size of the stimulus group (non-parametric Kruskal-Wallis test after confirming non-normality of data with Kolmogorov-Smirnov test, $F(5, 62) = 14.54, p = .012$). The probability of looking up reached near certainty when only three or more virtual characters nearest to the participant looked up. With only one character, the probability of looking up was more than 50%. Beyond three or more characters, all participants looked up.

![Figure 5.3](image)

**Figure 5.3**: The probability of participants who looked up depended on the number of virtual characters who looked up. Gray line indicates the percentage of participants across all trials who looked up during the 60 seconds prior to the instant when the characters looked up. Error bars indicate standard error of the mean.

There was a specific question in the survey which asked participants how many characters they noticed looking up (Q9). We used participants’ answers to this question to delve into
the probability of looking up. Fig. 5.4 shows average participants’ rate to this question as number of stimulus increases. Participants’ scores to this question were not significantly different for various stimulus group sizes (Kruskal-Wallis test, $F(5, 62) = 7.71, p = .17$).

![Figure 5.4: Average of participants’ rate to Question 9. Error bars indicate standard error of the mean.](image)

Figure 5.4: Average of participants’ rate to Question 9. Error bars indicate standard error of the mean.

Considering the reported inaccuracies in HTC Vive’s measurements in [17], especially the pitch measurements which are critical in our analysis, a sensitivity analysis was done to make sure that the presented results are valid through the possible errors. The results for a variation of ± 2 degrees in the head pitch rotation threshold are presented in Table 5.2.
Table 5.2: Probability of looking up sensitivity to the threshold

<table>
<thead>
<tr>
<th>Threshold (deg)</th>
<th>p value</th>
<th>F(5,62)</th>
</tr>
</thead>
<tbody>
<tr>
<td>-15</td>
<td>0.0124</td>
<td>14.54</td>
</tr>
<tr>
<td>-17</td>
<td>0.0125</td>
<td>14.54</td>
</tr>
<tr>
<td>-19</td>
<td>0.0125</td>
<td>14.54</td>
</tr>
</tbody>
</table>

**Proportion of time looking up:** The proportion of time after the stimulus is triggered that is spent by the participants looking up (Fig. 5.5) depended on the size of the stimulus group (non-parametric Kruskal-Wallis test after confirming non-normality of data with Kolmogorov-Smirnov test, $F(5,62) = 29.89, p < .0001$). This result was not sensitive to the variation of $\pm 2$ degrees in the threshold as well (Table 5.3).

![Figure 5.5](image)

Figure 5.5: Proportion of time that participants looked up depended on the size of the stimulus group, with more time spent as the number of virtual characters who looked up rose. Error bars indicate standard error of the mean. Post-hoc tests with Bonferroni correction indicated that the following pairs were significantly different: (1,10), (1,15), and (2,10).

Post-hoc tests with Bonferroni correction revealed that time spent looking up by participants when the size of the stimulus group was 10 was significantly more than when the size
of stimulus group was 2 or less. Time spent looking up with 15 characters was significantly more than time spent looking up with 1 character.

Table 5.3: Proportion of time looking up sensitivity to the threshold

<table>
<thead>
<tr>
<th>Threshold (deg)</th>
<th>p value</th>
<th>F(5,62)</th>
</tr>
</thead>
<tbody>
<tr>
<td>-15</td>
<td>&lt; 0.0001</td>
<td>28.61</td>
</tr>
<tr>
<td>-17</td>
<td>&lt; 0.0001</td>
<td>29.89</td>
</tr>
<tr>
<td>-19</td>
<td>&lt; 0.0001</td>
<td>29.97</td>
</tr>
</tbody>
</table>

Comparison with real-world experiments: Compared to the real-world experiments [20, 21], the probability of looking up in a virtual environment grew faster and all the participants who were in trials with three or more stimulus looked up (Fig. 5.3). Results of this experiment were also similar to real-world studies in terms of the time that the participants spent in looking up [21], as participants spent significantly more time looking up when the stimulus group was larger.

There are some differences between real-world experiments and this study that could contribute to this, as follows:

- First, the stimulus group was chosen based on their proximity to the participant in this study and proximity has been shown to play a significant role in determining the consistency of response [21]; also, participant’s view was generally clear and not occluded.

- Second, the head-mounted display field of view is 110 degrees and smaller than the human visual field [31]; this would eliminate other distracting cues in the environment.

- Third, most of the participants did not have experience with VR, so they were particularly attentive to everything inside the VR.
Results of this experiment were also similar to real-world studies in terms of the time that the participants spent in looking up [21], as participants spent significantly more time looking up when the stimulus group was larger.

Finally, a contagion model proposed in [21] states that the proportion of crowd looking up as a function of the stimulus size $N$ is

$$P(N) = m \frac{N^k}{(T^k + N^k)}$$  \hspace{1cm} (5.1)

where $m$ is the maximum proportion of the crowd that will look up, $T$ is the stimulus size at which half the crowd will look up, and $k$ controls the rate at which the stimulus group size determines the level of contagion.

We fitted this model to our data by minimizing the sum of square error between values predicted by this model and those obtained in the experiment to find values in our case. The obtained values and other studies’ values are presented in Table 5.4.

<table>
<thead>
<tr>
<th>Study</th>
<th>m</th>
<th>T</th>
<th>k</th>
</tr>
</thead>
<tbody>
<tr>
<td>Milgram et al. [20]</td>
<td>0.92</td>
<td>1.2</td>
<td>1.05</td>
</tr>
<tr>
<td>Gallup et al. [21]</td>
<td>0.66</td>
<td>7</td>
<td>1.38</td>
</tr>
<tr>
<td>Our study</td>
<td>1</td>
<td>0.7</td>
<td>1.7</td>
</tr>
</tbody>
</table>

These variables characterize the response type. The maximum probability of looking up ($m$) in our study is pretty close to the real-world experiment [20] that we replicated in the virtual environment. The lower $T$ value shows that in the virtual environment it needed considerably less stimulus number for the participants to look up. In addition, considering that $k > 1$ indicates a quorum-like response, participants’ responses in the virtual environment saturate sooner.
5.3 Conclusion

**Virtual environment simulation**: Experiments with large groups of people are difficult to conduct. Virtual reality presents a unique opportunity to study human behavior in crowded settings provided the response to virtual crowds resembles that in the real world [3]. Therefore, the first challenging part of our study was a realistic and high-performance virtual environment simulation. To do so, we needed to make sure that all aspects of the simulation were consistent with real-world experiences.

In this context, all appearance details including the geometry and arrangement of the virtual environment, humanoid characters, animations and transition between them were modeled similar to real world. In addition, a combination of two crowd models were used to guarantee realistic collision-free movements. The characters’ turn rate and stepping speed were synchronized with their velocity to avoid any sudden movements.

**Hardware and software restrictions**: We intended to design a larger and more crowded environment and started to develop a larger walkable virtual reality platform, but it was not practical due to the huge delay that we ended up with. We replaced our platform by HTC Vive setup that provided a limited walkable area for the participant. On the other hand, this limited walkable area that was equivalent of our experimental arena was a bottle neck for modeling a larger crowd. As the number of virtual characters increased, their density and simultaneously the probability of collision increased. It needed a smaller time step for the crowd model to guarantee collision-free performance, which was adding huge delay to the simulation.

Considering the aforementioned technical limitations, we used the current dimensions and crowd level in the virtual environment. It should be mentioned that even in the current setup we had to significantly optimize the script and modify the crowd model to keep the processing
time at the lowest level. Moreover, interaction forces between virtual characters and also virtual characters and stationary objects in the environment were ignored if their relative distance was more than 4 meters, data collection was reduced to the virtual characters that were within 4 meters from the participant, and data collected every 10 frames in addition to optimizations in Unity interface itself.

**Reproducing social influence in crowds:** The results of this study showed that virtual reality could be used as a viable platform in human behavior studies. Specifically, the participants perceived the virtual environment as realistic and looked up with increasing probability as the stimulus group size grew. The virtual environment successfully triggered behavioral contagion; it can be customized to test hypotheses in crowd behavior and for treating social disorders [32].

**Assessing and bridging the gap between virtual and real environments:** Comparison between the conducted experiment in the developed virtual environment and real-world experiments showed that participants’ responses would not be exactly same to the virtual and real stimulus. There are methods in the literature that quantify such gaps, including presence questionnaires, which we used to inform our own questionnaire; participant movement, which should be realistic and free of unseen constraints; and postural responses, which should be as if real [33, 34]. In addition, the type of virtual reality platform might affect participants assessment of the environment in some areas, including rehabilitation studies [35].
REFERENCES


[34] Slater, M. Place Illusion and Plausibility Can Lead to Realistic Behavior in Immersive Virtual Environments. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 364(1535), 3549-3557, 2009.