Crowd Simulation and Virtual Reality Experiments for 2010 Love Parade Disaster

Master Thesis

Author(s):
Zhao, Hantao

ETH

Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Crowd Simulation and Virtual Reality Experiments for 2010 Love Parade Disaster

Master Thesis

Hantao Zhao

10 December, 2016

Advisor: Prof. Dr. rer. nat. Dr. h. c. Dirk Helbing

Department of Computer Science, ETH Zürich

Abstract

Highly dense crowds have been a major threat for large festival event in public space. The 2010 Love Parade disaster provides large amount of data, such as research papers, professional reports, and video footages that can be used to research this disaster from different perspectives. In this study, we aim to provide a computer aided crowd disaster simulation tool that can be used to help to analyze the disaster at an organizational level. From the simulation results, administrative methods can be suggested and tested in order to prevent the disaster. However the Love Parade simulation from previous work of Wolff faces the challenge of unrealistically high crowd densities. The first step of this study is to improve the model and the navigation mechanism so that the simulated crowds densities remain realistic. By applying the *SocialForceModel* for steering agents and testing different crowd management strategies, we aim to replicate the disaster in order to manage crowds in a virtual environment. Once the simulated crowd was sufficiently similar to the real crowd, empirical studies were executed to verify the simulation results. A state of the art Virtual Reality(VR) device is used to generate an immersive experience of critical parts of that event. The observation of participants' emotional arousals and stress levels would help decision makers to better understand crowd management strategies. In the end we verified the hypothesis that better crowd management strategy can help to reduce the stress level from a first person perspective simulation. Using a simulation tool based on the Social Force Model and Unity game engine, we provided a sophisticated framework to examine crowd management strategies.

Acknowledgements

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Crowd disasters caused by high density areas during events such as festivals or evacuations are a primary concern of sociology and computer science researchers. Their chaotic and unpredictable character makes them a challenging yet interesting domain. Simulations by computer programs aim to address this challenge using both realistic replay and strategic prevention method. A better understanding of crowd management can help to avoid similar disasters in the future. It is relatively hard to project the crowds movement to a field experiment in reality, but computer simulation allows it approachable and feasible in a virtual way. This study used computer science approaches to investigate how the crowd disaster at the 2010 Love Parade electronic dance music festival in Duisburg of Germany happened and what could have been done to prevent it.

In the first part of this study, we recreated the Love Parade disaster using computer simulations with the Unity game engine. A framework was built to generate and steer the crowds, as well to collect the data during the simulation. All of the simulations are built in 3D VR environment in order to provide veracious visualization of the crowds and the Love Parade field. Next the crowd data were analyzed to check the usability of simulation results for a first person perspective experiment. This process helps us to better understand crowd formation and movement. Subsequently different management strategies were deployed to the environment and crowds to test their hypothesized effects. Based on multiple parameters of the crowds we want to determine the optimal strategy to prevent the disaster.

Furthermore an empirical study was designed to test the proposed strategy in another dimension. We employed the simulation results into a first person perspective replay using a VR device. Crowd panic is one of the causes for crowd disasters since it leads to tension and increased aggressiveness [6]. Panic effect are seldom taken into considerations in crowd simulation. A simulation replay mechanism that enables projection of the crowds flow towards

first person view can help to demonstrate the simulation results so that researchers can use it for empirical study. The anticipation from an empirical study is to test whether simulations of the crowd disaster are sufficiently real to induce changes in the observers' physiological states. The analysis of participants' reaction to the crowd simulation can provide another layer of complementary information to support a management strategy.

Wolff's[39] work on crowd simulations of Love Parade is foundational to this study. In the initial phase of this project, a crowd generation mechanism and 3D model environment of the event have been extracted from Wolff's work. One challenge of the prior simulation is its unrealistically high densities. The *Social Force Model*[11] is one possible solution for this issue, because it increases the repulsive forces between the agents. Moreover Helbing and Mukerji suggested a variety of organizational changes to avoid the disaster[12]. But it has not yet been demonstrated with empirical evidence. Simulations of enumerated strategies may serve the role of justification. To further test the simulation results, a first person replay will be deployed based on the generated results. On the other hand studies have depicted that crowdedness could correlate with stress level[27] and it matches our understanding and prediction about the correlations between crowdedness and crowds panic.

In order to capture the participants level of stress, physiological device was utilized during the empirical experiment. In addition subjective stress level rating questionnaires was used to provide one more measure of nervousness. The purpose of the empirical study is to find a reliable method to verify the performance of different management strategies by using first person replay and stress level assessment. We expect to find evidence that a better crowd management strategy can produce a lower level of stress. Combined with findings from simulations and empirical studies, we will be able to verify potential prevention strategies for Love Parade disaster.

In general with this study, we aim to provide a complete crowd simulation and management strategy verification mechanism to enhance decision making for large events. This framework could allow future event managers to easily test their event environment using computer simulations. For local small scale events it could provide a prevention mechanism to avoid crowd disasters such as 2010 Love Parade. Eventually this framework may provide a platform for crowd management and forecasts into the future. Large event manager of public space(e.g. train station, airport, city festival) or theme parks (e.g. Disney Land) can benefit from the simulation results to manage the crowds by altering the pedestrian flow or environment in evacuation scenarios on highly crowded events. Together with crowds monitoring information, the system can ultimately predict accidental event and provide prevention methods before it would happen.

Chapter 2

Related works

Love Parade disaster Love Parade as a popular music festival in German was firstly organized in 1989. The Love Parade disaster took place in Duisburg on July 24, 2010[39]. Helbing and Mukerji's detailed analysis of the festival area pointed out the flaws of the organizational plan: inconvenient maximum flow capacity and potential safety issue caused by the use of tunnel[12]. Their analysis presented possibilities to prevent the problems that contributed to the disaster. Except for administrative decisions and better communication protocols between the organizers, the following suggestions were proposed related to crowd management strategies: 1. a different flow concept such as separation of the inflow and outflow; 2. remove the obstacles on the ramp; 3. usage of the side ramp; 4. redirection methods can be applied by the police cordons once it gets too crowded[12]. It has also been studied the liaison between high density and crowd disaster. In high density areas overcritical turbulent and pressures can trigger the falling and trembling[10].

Besides Love Parade in 2010, other disasters are also focuses of sociologists. Research has used videos as data source from crowd disasters to evaluate computer algorithms (e.g. the study about the disaster during a Hajj event in Mina[10]) . Krausz *et al.* used a system for flow computation to avoid the vision detection of the camera, which is usually difficult and challenging due to individual tracking and visual occlusion[20]. Pretorius *et al.* used buildingEXODUS model to simulate the Love Parade disaster and some of the proposed strategies could eventually provide a solution for the highly dense crowd[32]. Reuter *et al.* provided another perspective of how modern social media manages criss events for instance in Love Parade disaster, which allows citizens to participate and use mobile phones or other media as infrastructures to influence the event lively[33].

Social Force Model   Initially proposed in 1995, *Social Force Model* has demonstrated its use for crowd steering and pedestrian simulation. By applying fluid dynamic traffic concept and microscopic modeling, it serves the role of describing self organized interacting pedestrians in a realistic way[8] [11]. Its utilization in disaster simulations is an useful pattern for systematic studies of panic behaviors, which usually involves a mass scale of crowds and collective phenomenon for crowd dynamics[9] . Driven from the concept of Newtonian law of motion, the model follows the nature of physical laws such as velocity and acceleration. Like any other kind of similar models, the tuning for its parameters requires a lot time and effort for analysis work. Johansson *et al.* provided an advanced algorithm to determine optimal parameter specifications for the Social Force Model[16]. Widely studied and applied, Social Force Model has driven a great amount of interests from researchers. Mehran *et al.* suggested a crowd behavior detection method which captures the abnormal behaviors without tracking specific individuals[26]. Moussaïd extended the model by applying it to social groups studies. The model well predicts the walking patterns extracted from video footage by analyzing the organizations of social groups[30]. Social Force Model has been widely used in sociology research due to its simplicity and usability as physics based approaches for crowd simulation. Yet Koster *et al.* argued that the differential equation is actually non differentiable and discontinuous at critical points, which would cause the loss of accuracy in numerical approximation. Instead he suggested a Modified Social Force Model to remove discontinuities and at the same time increase the computation speed[19]. Furthermore he warned that the choices of parameters could contribute to the phenomenon of oscillations if they were not chosen carefully[21]. Thus he suggested to use the elliptical specification model in Johansson *et al.* 's paper regarding specification of the social force pedestrian model [16] or adding the forces to avoid the risk[21].

Crowd simulation   In order to replay a realistic disaster scene, a crowd simulation mechanism is essential to this study. SteerSuite[34] is an open framework that allows user to simulate steering algorithms and crowds behaviors easily by providing tools for facilitating, benchmarking and testing. It can provide the core of agent navigation rationale, which is critical for a successful crowd simulation. Moussaïd *et al.* [28] used a heuristics-based model to study high-density situations during disaster. This approach can be applied to suggest environmental setting improvement for mass events in order to avoid crowd disasters, which is the same purpose of our study. Singh *et al.* demonstrated that a single platform which combines a variety of steering techniques could perform robustly and efficiently[35]. Crowd simulation done by Wolff's analysis of the influence of the environment and crowd parameters on the 2010 Love Parade disaster gives this study a platform to

start with[39].

Virtual Reality   Due to its nature of complexity and difficulty, crowd simulation visualization has remained in a relatively elementary level for a long time. 2D graphics dots and lines were commonly used. Massive crowd movement requires high computational power, which wasn't easily accessible for researchers decades ago. Nowadays computer power has profoundly evolved to allow feasible 3D simulations even on PC. The benefits of more immersive feeling and 3D graphic presentation extend the possibilities for experimental design. Game engines such as Unity 3D facilitate Virtual Reality development in a lot of levels such as graphic rendering and animations. State of the art Virtual Reality device Oculus provides relatively cheap and accessible technology for research purposes[4]. As an open source framework, *Experiments in Virtual Environments*(EVE) allows researchers to deploy their VR experiment in a systematic way [15]. Combined with its functionalities, an experimental protocol can be established in Unity easily. Kallman *et al.* reviewed recent development of real time planning for navigation in virtual world[17]. His analysis in applications of multi agent simulation in virtual world gave a profound overview of various algorithms for path searching. Other study has been using Virtual Reality technology for accidents scene replay, such as the study of tunnel driving accidents and behavioral training [18].

Crowds density and stress level   One of our research goals is to study the correlations between crowd management strategy and individual's stress level. The question of if being crowded would make people feel more nervous than not is essential for us. Freedman *et al.* gave a detailed investigation of crowding phenomenon in a systematic way by providing psychological analysis of human and animals behaviours[6]. Broadening similar concept to VR environment would answer our question in a way. Moussaïd *et al.* demonstrated that crowds in VR controlled by human subjects follows similar patterns of real crowds as observed in real-life crowded situations[29]. It has been illustrated that virtual environment can indeed provide an immersive feelings of being interacting with crowds that behaves in a plausible way[31]. Merriman *et al.* stated that being crowded resulted in a cost in performance to the spatial tasks, compared with the non crowded group[27] .

The stress level of an individual participant's one of the primary focus of crowd management simulation. A subjective evaluation of stress level can provide a record on participant's emotional change before and after the task. Mathews *et al.* developed Dundee Stress State Questionnaire model to measure this change[24]. Although the Worry state provides more cognitive aspect of the participant[13], Short Stress State Questionnaire(SSSQ) has been

presented to be able to apply as the metric in stress state assessment, in three dimensions: Engagement, Distress and Worry[3]. All of the three factors are highly related with crowds behavior and stress level assessment. In the form of a pre questionnaire and post questionnaire, SSSQ can be utilized to observe the increase or decrease of the stress level of the participants.

Electrodermal activity(EDA) analysis is widely used as an analyze tool in judgment and decision making research. The monitoring of the physiological process can reflect the attention, habituation, arousal and cognitive effort[5]. Skin conductance, as one of the forms of EDA, can support a relatively simple and cheap approach for EDA analysis. Benedek *et al.* points out that skin conductance response(SCR) amplitude can used to reflect sympathetic activity, in particular the phasic sympathetic activity[1].

Chapter 3

Love Parade crowd simulation

The first step of our work is to establish a crowd simulation framework. This would allow us to freely choose crowds generation, steering and environment manipulation methods. This chapter covers the structure of the crowd simulation mechanism and the rationales of the projection of the disaster.

3.1   Love Parade disaster overview

The Love Parade in 2010 was located in a previous freight station in Duisburg. In Figure 3.1 an illustration of the festival area is presented with detailed environment setting and event related data. The two side of the tunnel was used as entrance to the festival area. A main ramp, a 26 meters wide inclined road, was in charge of leading the visitors to the festival area. The ramp was surrounded by fences. During the event four police cordons were established to control the flow, starting from 15:50. Between 15:30 to 16:00 the congestion around the main ramp began to form. There were people trying to exit the flow from fences and stair. The fences in the main ramp area were removed around 16:20, due to the severity of the condition[39].

3.2   Prior work

The crowd simulation was based on Katja Wolff's thesis[39]. During that study a crowd simulation mechanism in context of Love Parade disaster was built with Unity game engine platform. The fundamental crowds generation and steering strategy were done previously and kept during the present study. The data acquisition from video footage and online statistics was used to determine the crowds inflow and outflow in a way that was similar to the real event. The festival area was rebuilt with a true-to-scale 3D model and integrated into the Unity engine. The ground part of the tunnel consisted of the main ramp, the tunnel and the fences. The surface of the model was
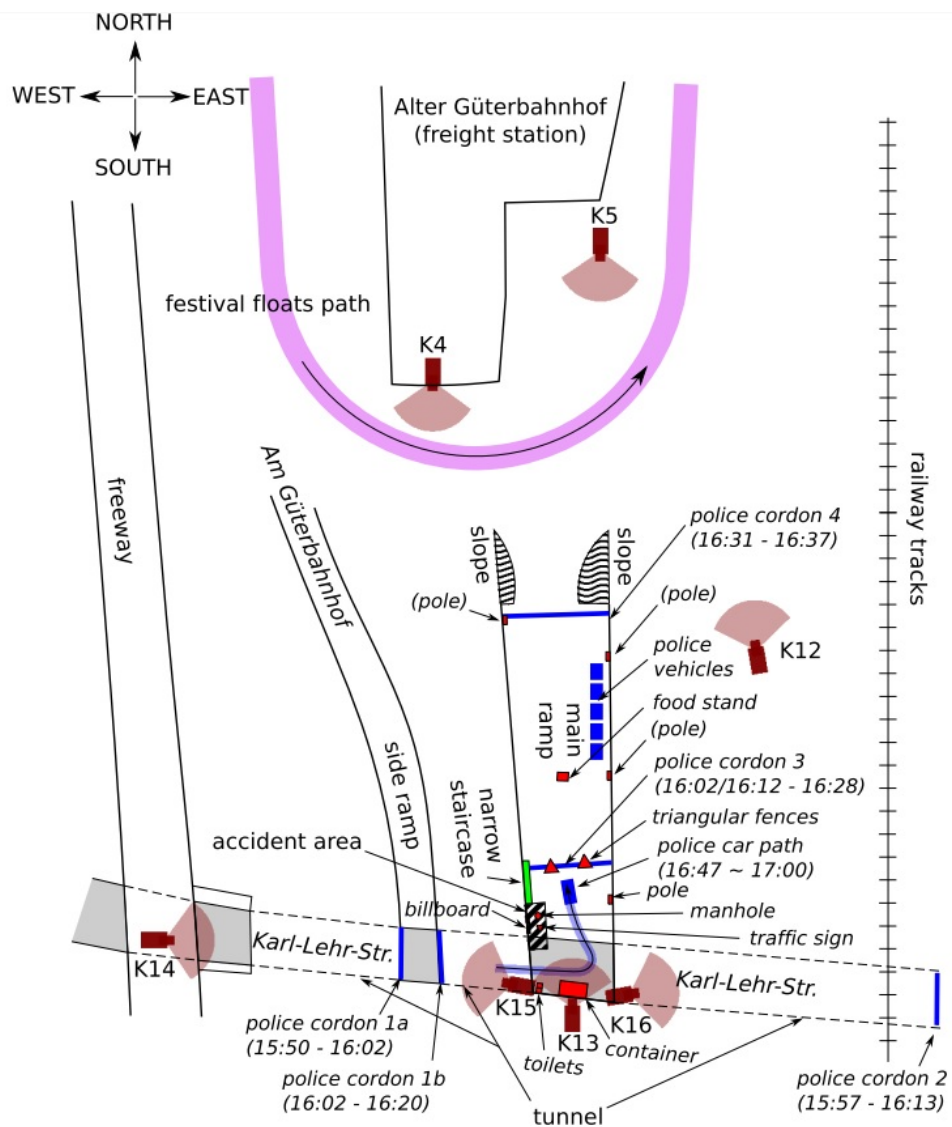
Figure 3.1: Illustration of the festival area
[12]

walkable for the simulated agents. The festival area was represented with a building for simplification.

Crowd generation and simulation The crowd generation needs to be conducted in a realistic way so that the simulated density number is similar to the density of the real event. We used the inflow and outflow statistics that were estimated by volunteers watching surveillance footage of the event[39]. Inflow and outflow numbers vary with time as the event progresses. A

linear interpolation between the number of people in the crowds and their velocity provided a distribution of crowds movement velocity. This has been done using the fundamental diagram proposed by Weidmann[37]. Wolff's applied the function with density expressed as

$$\rho = \mathcal{Q}/\mathcal{V} \tag{3.1}$$

where $\mathcal{Q}$ is the combined in- and outflow per second and meter, $\mathcal{V}$ is the visitors' velocity[39].

$$\mathcal{V} = 1.34 \times (1 - e^{-1.913.(1/\rho)-1/5.4}) \tag{3.2}$$

As shown in equation 3.2, the velocity varies with the flow number. Our primary focus is the time when the disaster happened, which was when the flow number was static. Thus the crowds velocity is constant at 0.702 m/s constantly after 15:20 on.
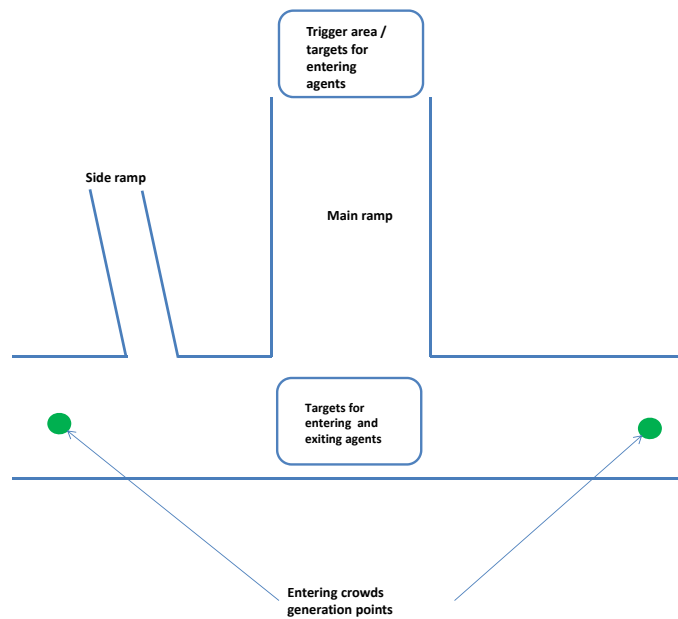
Figure 3.2: Overview of the targets and trigger areas on the ramp to navigate agents.

Shown in Figure 3.2, the crowds were generated in the green areas and once they reached their corresponded targets such as the end of ramp or the end of the tunnel, they were destroyed in order to save computational abilities.

Police cordons are simulated by a specific area where allows agents to halt or walk, corresponding to when the cordons are active and when they are deactivated[39].

The agents' controlling and generating system are consisted of the following essential scripts. Their connections are illustrated in Figure 3.3.
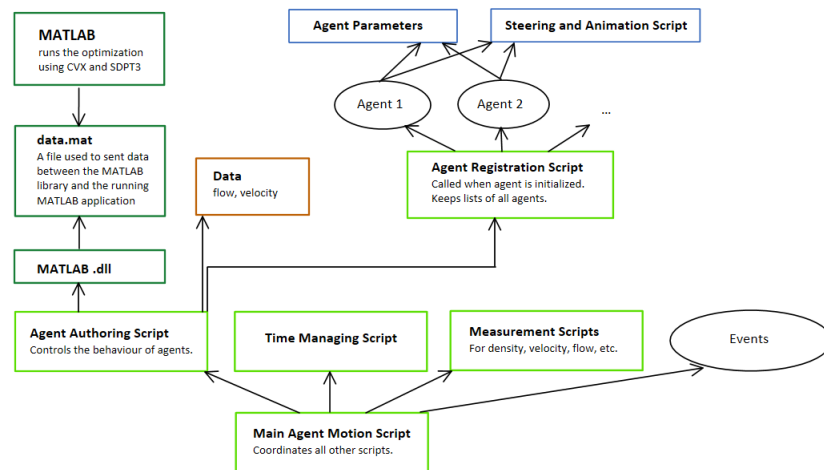
Figure 3.3: An overview of the Unity project structure
[39]

- **AgentParameters** contains agents information such as ID, velocity and target

- **Steering and Animation Script** takes care of navigation and animation mechanism

- **AgentRegisteration** initialization and registration of agents

- **AgentAuthoring** applies precomputed parameters on the agents(e.g. velocity and targets) and controls agents behaviors

- **Measurement Script** in charge of measuring and recording simulation results

- **TimeManaging** timing system of the simulation. Synchronizing the simulating elapsed time into the real world time scale.

- **MATLAB files** simulates the festival area agent behaviors after they enter into the festival area. Behaviors include getting food, floating around the main building, visiting the video wall and so forth. It is not a primary focus of this study.

Avatars In the simulation part the agents were represented with cylinders without complex animations in order to save computational power. Avatars with animations for the replay experiment will be presented next in Chapter 5.

3.3 Social Force Model

Social Force Model introduction One of the legacy issue from previous work was the unrealistic high density. The forces between agents were not well projected in to the simulation. Thus it is reasonable to choose a model that can well reflect the physic laws between the agents. Helbing *et al.* provides us a sophisticated pedestrian simulation model *SocialForceModel* that suits well to our study[8, 11].Systematic influences such as target, obstacles and other pedestrians form a complex force to each pedestrian. The force is formed by an acceleration force $\vec{f}_\alpha^0(\vec{v}_\alpha)$, a repulsive effects $\vec{f}_{\alpha B}(\vec{r}_\alpha)$ caused by obstacles and boundaries, and a repulsive interactions $\vec{f}_{\alpha\beta}(\vec{r}_\alpha, \vec{v}_\alpha, \vec{r}_\beta, \vec{v}_\beta)$between the pedestrians[8]. Shown in equation 3.3,

$$f = \vec{f}_\alpha^0(\vec{v}_\alpha) + \vec{f}_{\alpha B}(\vec{r}_\alpha) + \delta * \sum_{\beta(\neq\alpha)} \vec{f}_{\alpha\beta}(\vec{r}_\alpha, \vec{v}_\alpha, \vec{r}_\beta, \vec{v}_\beta) + \sum_i \vec{f}_{\alpha i}(\vec{r}_\alpha, \vec{r}_i, t) \qquad (3.3)$$

δ represents the importance factor of repulsive iterations. i is the assembly of the obstacles and β is the assembly of agents.

The acceleration force $\vec{f}_\alpha^0(\vec{v}_\alpha)$ is defined by the direction of the next destination e-alpha, desired speed v_α^0 and the current speed $\vec{v}_\alpha$:

$$\vec{f}_\alpha^0(\vec{v}_\alpha) = \frac{1}{\tau}(v_\alpha^0 \vec{e}_\alpha - \vec{v}_\alpha) \qquad (3.4)$$

Other obstacles, in this study the fences and boundaries such as walls together defines repulsive effects. In the equation 3.5 $\vec{r}_\alpha - \vec{r}_B^\alpha$ is the distance between the agent and the obstacles. V_B is a repulsive potential as a simple case.

$$\vec{f}_{\alpha B}(\vec{r}_\alpha) = -\nabla_{\vec{r}_\alpha} V_B(\|\vec{r}_\alpha - \vec{r}_B^\alpha\|) \qquad (3.5)$$

Forces between the pedestrians is defined as:

$$\vec{f}_{\alpha\beta} = A_\alpha^1 exp[(r_{\alpha\beta} - d_{\alpha\beta})/B_\alpha^1]\vec{n}_{\alpha\beta} * (\lambda_\alpha + (1 - \lambda_\alpha))\frac{1 + \cos(\varphi_{\alpha\beta})}{2}$$
$$+ A_\alpha^2 exp[(r_{\alpha\beta} - d_{\alpha\beta})/B_\alpha^2]\vec{n}_{\alpha\beta}$$

11

Aα represents respective interaction strength and *Bα* is the range of the repulsive interaction. Once the force is calculated it is converted into velocity and then normalized with the maximal speed. Finally the velocity is transferred by dividing the ratio between the desired FPS and current FPS[39]:

$$V = V/TimeManager.FPS * currentFPS \tag{3.6}$$

Time dependent parameters are not taken into account during this calculation since the time scale is relatively large in this simulation. Long waiting time might lead to a high acceleration force which could lead to unrealistically high density.

Parameter determination   The present parameters in the *SocialForcemodel* are individual parameters and partly dependent on cultural conventions[8]. In our study, we applied the parameter values provided by Helbing *et al.*'s original paper and the simulation result has an improvement over previous work. Thus following parameters are chosen by default in Table 3.1:

Table 3.1: SFM Parameters

Parameters		
Name	Type	Value
A_α^1	float	0
A_α^2	float	3
B_α^1	float	0.2
B_α^2	float	0.2
λ	float	0.75
δ	float	0.015

The other parameters such as neighborhood range and the importance of agent repulsive force are tuned during the simulation in order to reach a legitimate density level and crowds formation. They remain constant during the whole simulation.

3.4   Steering strategy

Crowds generation and steering mechanism   The navigation of the agents to the desired targets is a core function of the simulation. The agents' path should follow the general pattern of real pedestrians' choices. In order to achieve this effect, a combined steering mechanism is applied that uses the Unity built-in Navmesh system and the *SocialForceModel*.

Unity Navmesh is utilized to generate a path with multiple milestones. Before the simulation a navigable mesh, which consists of triangles arranged

to represent the 3D object, was generated in advance. A navigable path was subtracted from a polygon, which is a collection of vertice, edges and faces from the mesh that connects the starting and ending position. The path makes use of the targets, obstacles and scene geometry to provoke a nearest polygon of the created mesh that would calculate through the polygon neighbors until the final target is reached. This mechanism used $A*$ algorithm, also known as the best-first search, by finding the path that incurs the smallest cost[7]. Moreover we need to adapt each agent's speed based on the simulation time scale using formula 3.6, since a real time simulation with large amount of agents is not achievable and time consistency needs to be taken into account. The rendering frame rate is fixed at 25 FPS so that the rendering time differences can be converted in a stable way. *TimeManaging* class keeps updating the current framework FPS so that the agents can adapt their velocity based on that FPS number. The agent generation is processed in a similar way in order to make the spawning procedure frame dependent instead of real world time dependent.

Dynamic target   The first simulation results depicted that a singular Navmesh navigation strategy can lead to a result of partially high density area and incorrectly distributed path choices. Analysis of the density illustrates that the concentrated use of the corner that leads to the shortest path to the target is the primary cause of the over concentrated local areas, see Figure 3.4.
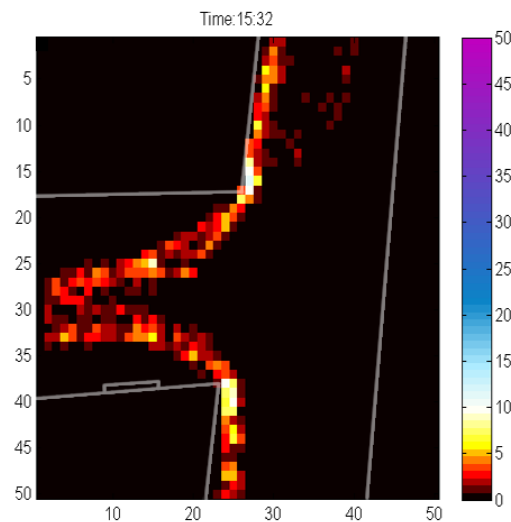

Figure 3.4: Original simulation density map at 15:32 from Wolff's prior work [39]

An approach called smart target was deployed to create a dynamic target

that allows the crowds to navigate between the milestones in a realistic way. The solution provides the agent a smart target that can be presented as an area. The area would allow the agent to flexibly enter and exit the target area.

Figure 3.5 illustrated an example of a steering process. A pedestrian enters the tunnel from left side and is firstly targeted at the center area of the crossing (Rectangle ABC). Naturally a pedestrian would choose the path that is close to point A, which would create a shortest path. In our approach, we created a probability tendency distribution (See Figure 3.6) to randomly assign the agents target between point A and point B. This tendency is generated from inflow observation of the camera data. In the video people tend to choose the shortest path, which is the connection between their starting position and the point A to enter the central area of the ramp. For example if an agent follows this mechanism, it will have almost zero probability to choose the point A or point B directly as the target. Yet it has a large probability (around 0.15) to choose the point that is close to point A. Since we want the path to have a realistic tendency, the probability of steering to point B will decrease. The sum of the probability for each agent in Figure 3.6 is 1. This pseudo randomization mechanism prevents a scenario in which all the agents choose one path and become too crowded in the corner. At the same time this mechanism is more realistic than using a total randomization betweens point A and B. Another factor that needs to be taken into account is the agents' starting positions. In Figure 3.5 it can be seen that Agent 1's distance to point A is closer than Agent 2. Instinctively Agent 1 will have larger chance than Agent 2 to choose the target that is near to point A. The steering probability would take this into account and choose the target for Agent 1 and 2 differently. In Figure 3.6 it is illustrated that Agent 2's probability has a shift towards point B compared to Agent 1, which corresponds to the initial position shift that is presented in Figure 3.5.

Similarly once the agents have reached their first targets, they will be assigned to the next targets and exit the area. Targets are the chosen between point A and point C. Instead of the previous example of edge AB, this time the crowds will have a tendency to choose the center path between point A and point C. In Figure 3.7 the probability of choosing the middle area between A and C is the largest. The probability of choosing targets that are close to A and C drops as the distance to these points decreases. Another target shifting happens between the left and right entering points from two sides of the tunnel. If an agent enters from the left side of the tunnel, then it has a tendency to choose a point close to point A, but if it enters from the right side, then a tendency towards point C is added. After the agent exits the area ABC they will be assigned to the next target in the end of the main ramp. A target assignment of edge DEF is considered in a similar way as edge AC, which has a tendency to have larger probability to navigate
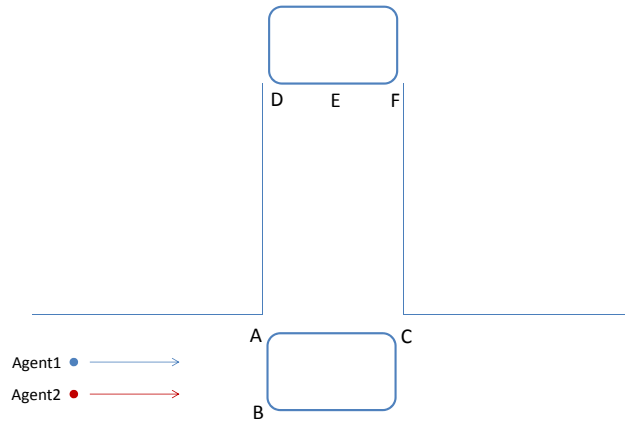
Figure 3.5: Steering process with smart target area in main ramp area

towards middle point E.

3.5 Implementation

Social Force Model algorithm   The core function of this implementation is the application of the Social Force Model. Algorithm 1 *SocialForceModel* describes the procedure of a model updates. Procedure Neighbor detection is independent from the procedure Controlled Update and only happens when *OnTriggerEnter* or *OnTriggerExit* are called. This procedure is responsible for recording the neighbor game objects of each agent. The neighbor detection range is a circle area with 1.5 meter diameters, located in the center of the agent.

The procedure Controlled Update is called once per frame update. It applied the Social Force Model to each agent in agent pool A. f_α^0 is the acceleration force. $f_{\alpha B}$ accumulates the avoidance force for obstacles and $f_{\alpha \beta}$ calculates the repulsive forces between the agents. The $v(f_\alpha)$ function converts forces to velocity and then the normalizes it according to the maximal velocity. After the calculation of the force, the position of each agent is updated respectively.

Data generation   Measurement data is saved as .txt files. The following statistics are recorded at each minute of the virtual time scale.
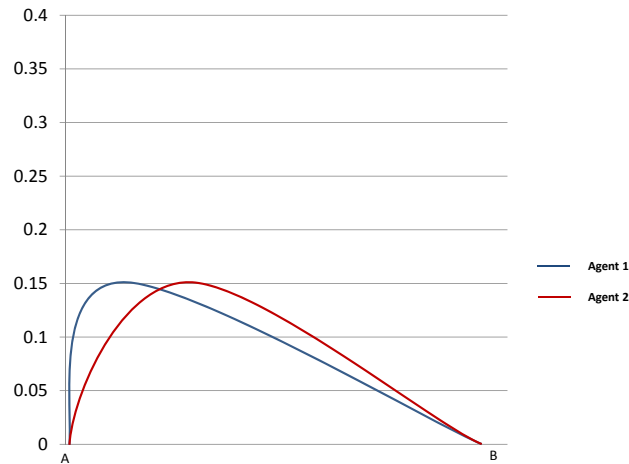
Figure 3.6: Steering probability to edge AB

- **Area density** measurement of the primary focus of the general ramp area. A 50x50 grid is created to continuously record how many agents are inside of each grid cell.

- **Flow measurement** measurement of the inflow and outflow at the center of the main ramp.

- **General frequency measurement** measurement of total agents that have been created, the amount of agents that has finished the whole route then been destroyed.

- **Congestion number** the number of agents that are in congestion status. An agent is defined as in congestion status if its movement distance is smaller than 1 meter after 60 frame updates.

- **Position data** the up to date position and initial starting position of each agent. Similarly this is recorded once per second of the virtual time scale.

Performance optimization methods The introduction of the Social Force Model requires an extremely large amount of computational power since detection and updates are called frequently. The neighbor range scale is a core factor because it decides the number of game objects that needs to be included in the SFM calculation. Increasing the neighbor range can lead to a smoothened trajectory and less oscillation effect. Yet this approach caused
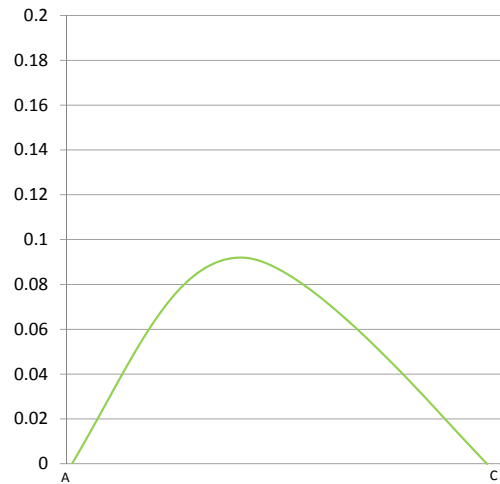
Figure 3.7: Steering probability to edge AC

an exponential growth in computational cost. With a compromise of a small neighbor range (a circle of 1.5 diameter), the total running time was reduced to around 48 hours.

Other optimization strategies were applied as well. Unity's profiler features the mechanism to list costly computations. Code redundancy and useless function have been removed according to the suggestions of profiler. Unity's light-mapping allows the engine to pre-compute the light map before the launch of the simulation. This approach saves the simulation from heavy real time light rendering. On the other hand occlusion culling disables the rendering of objects that can not be seen by the camera at that moment. Setting the static objects to *OccluderStatic* and executing the occlusion culling bake before running the simulation provide this feature. During the simulation avatars are replaced with simple cylinders so that the animation rendering can be avoided. In addition when the simulation is running, the program is in batch mode in order to minimize memory and CPU usage.

Algorithm 1 Social Force Model

---
procedure Neighbor detection
    if Object n *OnTriggerEnter* **then**
        $NeighbourCount \leftarrow NeighbourCount + 1$
        add n to $\mathcal{N}$
    end if
    if Object n *OnTriggerExit* **then**
        $NeighbourCount \leftarrow NeighbourCount - 1$
        remove n from $\mathcal{N}$
    end if
end procedure
procedure Controlled Update
    for each $a \in \mathcal{A}$ **do**
        $f_\alpha^0 \leftarrow v_\alpha^0 * e_\alpha - v_\alpha$
        for each $n \in \mathcal{N}(a)$ **do**
            if $type(n)$ = obstacles **then**
                $f_{\alpha B} \leftarrow f_{\alpha B} + f_{\alpha B}(n)$
            end if
            if $type(n)$ = agents **then**
                $f_{\alpha\beta} \leftarrow f_{\alpha\beta} + f_{\alpha\beta}(n)$
            end if
        end for
        $f_\alpha \leftarrow f_\alpha^0 + f_{\alpha B} + f_{\alpha\beta} * \gamma$
        $v_a \leftarrow v(f_\alpha)$
    end for
    for each $a \in \mathcal{A}$ **do**
        update agent a position
    end for
end procedure

Chapter 4

Love Parade management strategies and simulation results

The simulation framework described in Chapter 3 now allows us to assess combinations of viable management strategies and run crowd simulations based on them. As we already known from Chapter 2, Helbing and Mukerji already provided multiple optimal strategies prior to this study[12]. This chapter answers the question of whether this crowd simulation framework can be used as a reliable approach to test the strategies.

4.1 Management strategies

Management strategies were implemented and simulated by applying organizational changes such as crowds' generation locations or the removal of police cordons, independent of the original setting. The effectiveness of a strategic change can be examined according to multiple standards: general area density around the ramp, regional density at the accident area of the ramp, number of congestion crowds , casualties and so forth. If a strategy can reduce one or more than one of the numbers above, it can be considered *effective*. Another factor is the number of agents that finished their traces and then either entered the festival area or exited the tunnel. This number is defined as the finished number. In contrast to the other metrics, a strategy is effective if it increases the finished number.

Based on the analysis[12] and prior studies, we came up with following five hypotheses to be validated.

Hypothesis (H1): Removal of the police cordons is effective. □

Four police cordons are formed during the peak crowded hours of the event at different locations indicated in the Figure 3.1. The initial purpose was to control the inflow and outflow so that congestion could be avoided. Yet since

the inflow continuously grew from both side of the tunnel the police cordon actually increased the severity of the congestion. If the police cordons were not formed then the inflow can be guided to enter the festival area and reduce the number of inflow for later.

Hypothesis (H2): Earlier removal(before the beginning of the simulation) of main fences is effective. □

The main fences in the accident area were placed as the barrier of the entrance port. While the crowdedness continued to grow, the organizer decided to remove the main fence at 16:20. This procedure is applied to all of the simulations. Naturally we could presume that if the fences were removed earlier there might be a difference.

Hypothesis (H3): Utilization of the side ramp as entering and exiting channel is effective. □

Show in Figure 3.1 a side ramp is on the left side of the main ramp but was not utilized at all during the event. It could have been used as a second entering or exiting channel for the crowds.

Hypothesis (H4): Separating the inflow and outflow by using the side ramp as exit channel is effective. □

Followed by the previous hypothesis, if the organizer can separate the inflow and the outflow during the event then crowds congestion may not be formed at all. It can at least quicken the exiting and entering process by accelerating the crowds movement then reduce the crowds turbulence.

Hypothesis (H5): Applying the removal of the police cordons to the other three strategies gives an effective improvement. □

We also want to combine different strategies so that an enumeration of possibilities can be tested as well. The police cordons actually increased the level of crowdedness by slowing down the crowd movement. Considering the nature of the strategies, combinations of removal of the police cordons and the other three strategies are worth validating.

4.2 Simulation

4.2.1 Density analysis

Density is the number of agents per area range. It can be defined as $\rho = numberofagent/area$. There are two density levels that are interesting for us: the **regional density** at the accident area (Figure 3.1) and the **general density** around the main ramp area (Figure 4.1).

The **regional density** is defined by the inflow and outflow per second per meter in the accident area, computed by using the following formula of Weidmann [37]. In this case regional density is $\rho = numberofagent/area = n/A = (qin + qout)/V$. qin and $qout$ are the inflow and outflow numbers per unit of space and unit of time. V is the velocity. Total length of the accident area is valid if defined as the effective width of the ramp(L=10.59m). Yet we apply 26 meters as the the total width since agents are able to spread out after they passed the narrowest part of the ramp, of which its total width is necessary for calculation[39]. So the adapted calculation is $\rho = \dfrac{qin + qout}{L * V}$.
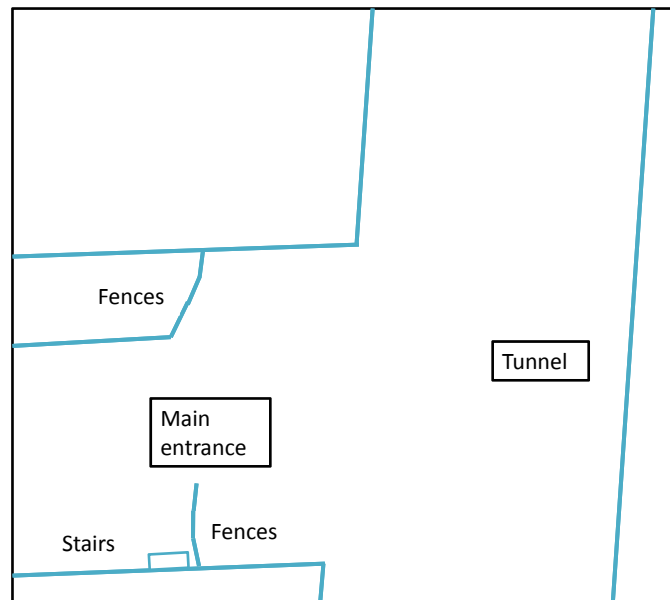

Figure 4.1: Density recording area. The areas within the black rectangle was constantly monitored.

The **general density** is the counting for agents in a 50X50 grid (See Figure 4.1). Both of the above densities have unit of ($agent/m^2$) and are recorded once per simulated minute.

4.2.2 Casualty calculation

In 2010 Love Parade, the disaster escalated when the people became too crowded and then suffocated and died. It is because "once the crowd density exceeds between 4 or 5 people per square meter, congestion can build up quickly, which implies high risks for people to stumble or fall (particularly if the ground is uneven). Therefore, injuries can easily happen[12]." Thus the concept of danger zone needs to be introduced. We define the density areas where the number of agents exceeds 5 per square meter as the danger zone(Danger zone$\geqslant$ 5). Similarly this concept can be introduced to danger zone that exceeds 7 people per square(Danger zone$\geqslant$ 7). We chose this number based on the pilot simulation run in which the zones with larger than 7 agents rarely happened but still existed. This number is generated from the regional density data of the whole ramp area. Furthermore the danger zone number can be used to calculate the number of casualties. Once a Danger Zone($\geqslant$ 5) is formed we estimate 1 people within the danger zone may die from suffocation and 3 people may be injured by falling or being stepped on. These numbers are only approximated to facilitate the estimation process.

4.2.3 Congestion

One characteristic of the simulation was the formation of congestions. When crowds formed a flow from three directions in the fence area, a congestion was easily formed. Figure 4.2 illustrates the moment when congestion is formed. The red cylinders were agents that were static. Once they recovered from congestion the color was changed back to green again. Altering the *SocialForceModel* parameters could reduce this congestion situation(e.g. reducing of the repulsive force between the agents or diminishing the range of the neighbor detection area). However if the number of agents is relatively high(in this simulation up to 70,000 agents) and there was a natural bottleneck formed by the fences, the congestion is almost inevitable. In addition if we wanted to project the simulation in a realistic way to avoid the extremely high density locations, the repulsive forces should not go below a certain level. Otherwise the density went unrealistically high again. Changing to another model or tolerance for high density may solve this problem. This remains future works so more details will be covered in the Chapter 8.

4.3 Results

The simulation results are presented in summary below. The simulation covered from 15:20 to 16:40 in virtual time scale of the event. Each simulation takes about 48 hours on PC. The strategy abbreviations are respectively:

OP Original simulation of the event.
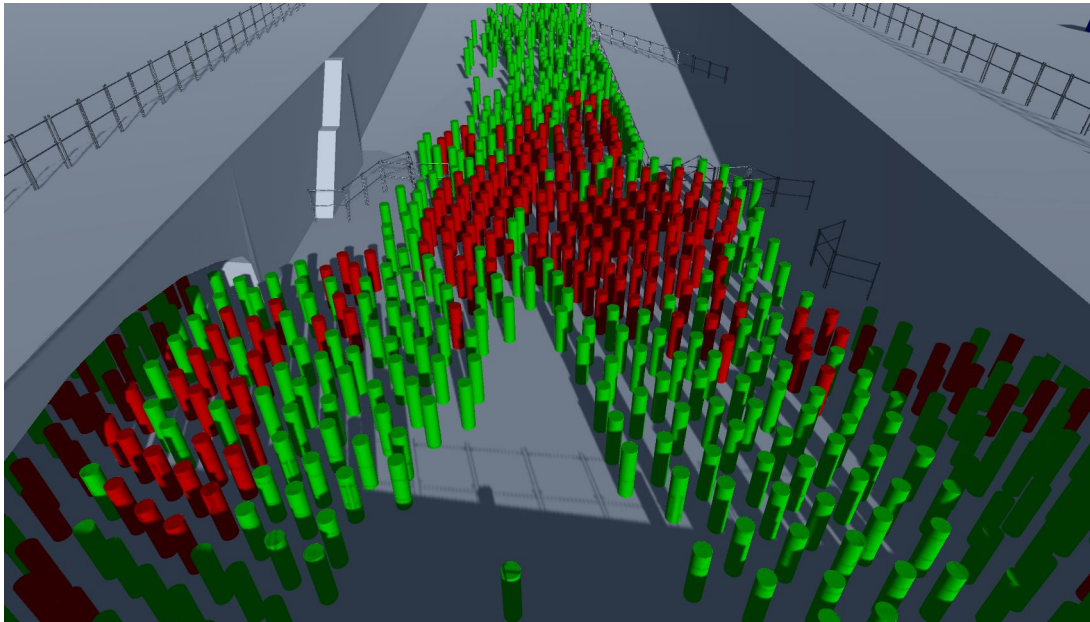
Figure 4.2: Formation of congestion

ONP Original simulation of the event, without police cordon.

FP Remove fence from main ramp area.

FNP Remove fence from main ramp area, without police cordon.

SP Separate the inflow and outflow using side ramp.

SNP Separate the inflow and outflow using side ramp, without police cordon.

RP Open side ramp for entrance and exit.

RNP Open side ramp for entrance and exit, without police cordon.

The general simulation result is listed in Table 4.1. 'MAX' represents the maximal density area($agent/m^2$) during the whole procedure. The danger zone counters are separated into two categories. The number of danger zones is the sum of danger zones over time(15:20-16:40). 'D($\geqslant 7$)' is the sum of danger zones that have more than or equal to 7 people per square meter and the 'D($\geqslant 5$)' is the sum of danger zones that has more than or equal to 5 people per square meter. 'Finished' represents the number of people that reached their final targets. 'ADen' represents the general density, which is the average number of people per square meter per minute($agent/m^2 \cdot minute$), in the 50X50 regional area of the ramp. 'ACon' represents the average congestion number per minute($agent/minute$) .

Table 4.1: General simulation results

Strategy	MAX	D($\geqslant 7$)	D($\geqslant 5$)	Finished	ADen	ACon
OP	12	39	356	17	0.904	34703
ONP	7	1	213	1	0.916	34721
FP	8	24	364	21	0.833	35249
FNP	8	12	400	11	0.852	35255
SP	3	0	0	10522	0.223	4278
SNP	2	0	0	13698	0.212	3270
RP	9	8	167	14	0.828	35760
RNP	6	0	134	39	0.8877	36172

The Social Force Model has reduced the unrealistically high density. Most of the MAX number of the simulations lies in a reasonable range, except for the OP group that has the MAX numbers that is slightly too high. In Helbing *et al.* 's simulation[10], MAX can go up to 10 so this simulation result is considerably reasonable. The average congestion number is still large in most of the groups except for SP and SNP. The danger zone number reflects the casualty number that is simulated. As illustrated unrealistically high density (> 10) is almost avoided and happened very infrequently. The number of agents that have finished the trace is low in general except for SP and SNP. This means that the congestion level is quite high in most of the simulations. The high number of ACon also supports this observation.

SNP and SP have enormous advantages over the other strategies. There are no danger zones in either group and the average density is only one fourth of the OP group. MAX is only 3 which suggests that there is no casualty in these two groups. There are a large number of agents that finish their routes and reach the final target. In the original simulation the removal of the police cordons has an effective improvement in MAX and Danger zone numbers. The removal of the main fence does not seem to have an effective change. RP and RNP both have an effective improvements compared to the Original group. They both reduced the danger zone numbers and improved the finish number. The ACon is higher compare to the original simulation but the average density is lower.

Furthermore the regional density in the accident area gives us an analysis of the focus area. Based on the calculation formula described in section 4.2, the time based regional density simulation results are illustrated in Figure 4.3. It is clear to tell that SP and SNP again have the lowest regional density level. In SP simulation the regional density is sometimes zero. This is caused by the blockage effect of police cordons. OP group has the highest peak density. In general the group without police cordon tends to have lower density than the same strategy with police cordon. Most of the groups have a density
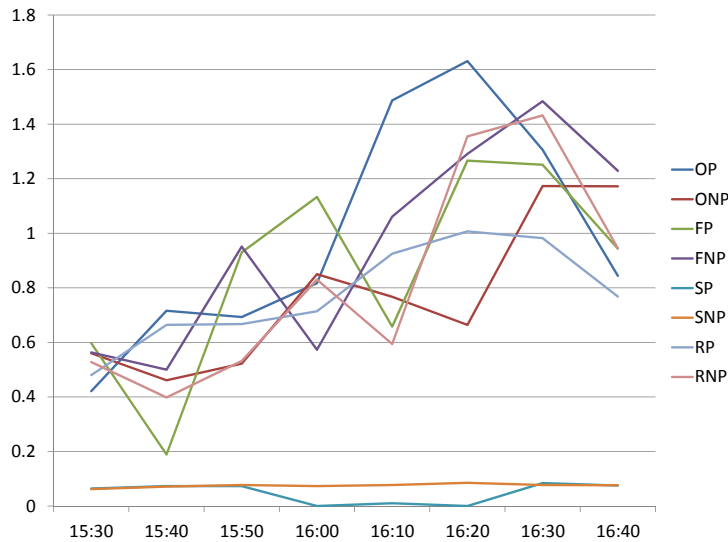
Figure 4.3: Regional density

drop after 16:20. This is caused by the remove of the main fence, which is a replication of the real event.

A heat map of density of regional area of ramp of all of the strategies at 16:00 is represented in Figure 4.4. SP and SNP groups have the lowest general density compared to the other groups.

A detailed plot representing Danger zone number $D(\geqslant 5)$ in the OP group is depicted in Figure 4.5. The danger zone number grows as time goes on and then decreases due to the removal of the main fence. The danger zone number was cumulative so this number can only give a general trend about the crowdedness. Yet the peak danger zone number still gives an approximate casualty number. The growth peaks at 27 danger zones around 16:20. If we estimate each danger zone to cause 1 death and 3 injuries as in previous section 4.2.2, the total casualty of this simulation is 27 death and 81 injuries. In real event there are 21 death and 500 injuries. With some approximation the simulation result is satisfactory.

4.4 Analysis

Based on this analysis, we can examine our primary hypotheses.

H1 removal of the police cordon is effective. From the comparison of OP and ONP, the numbers of MAX and danger zone drops after the police

ONP

OP

FNP

FP

RNP

RP

SNP

SP
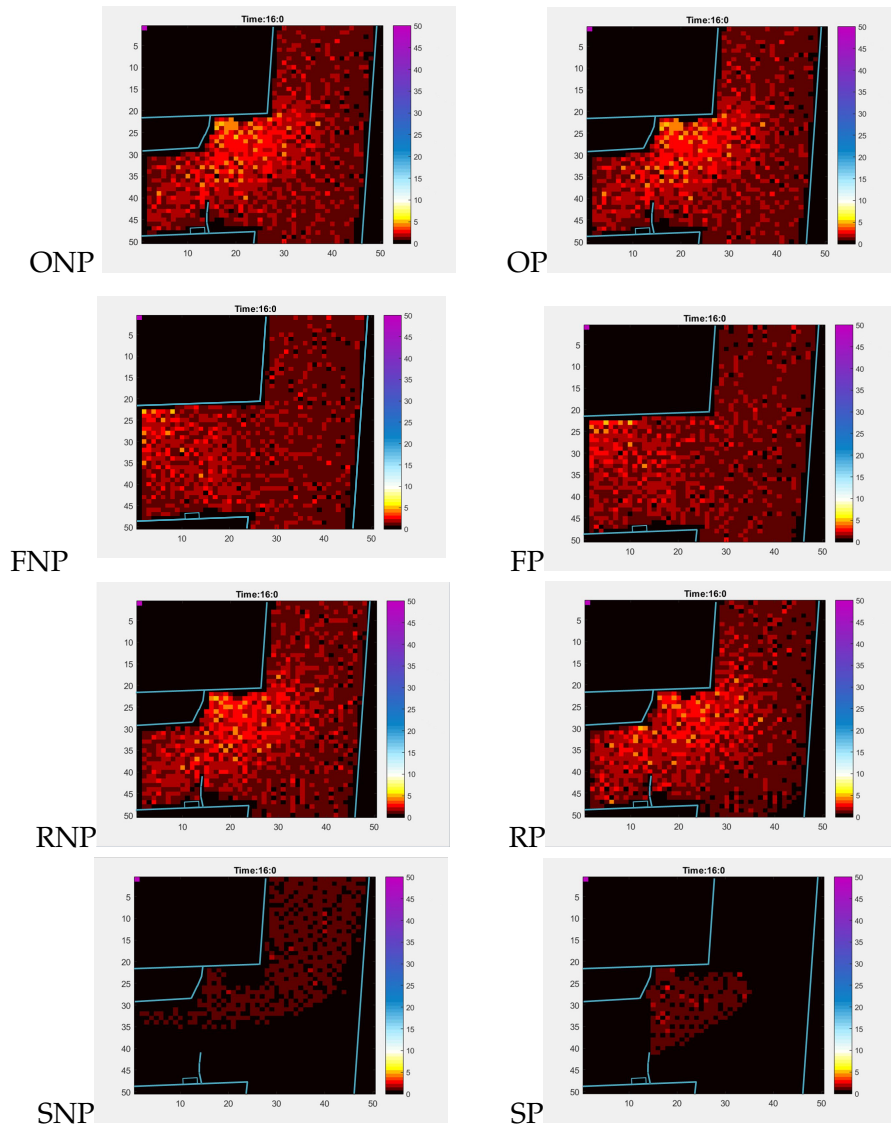
Figure 4.4: Heat map of general ramp area at 16:00
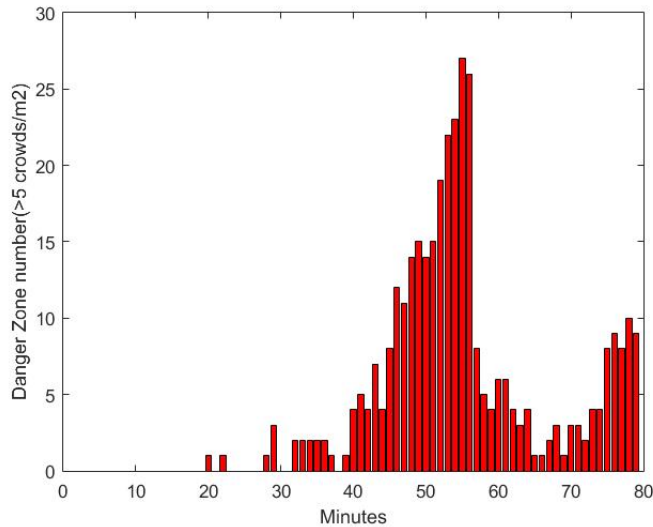
Figure 4.5: Danger Zone($\geqslant 5$) number change of OP

cordon is removed. The congestion numbers are both high and the relative difference is small. It is the same with finished number, where little number of agents can finish the routes. Thus these two numbers can not be used to judge which strategy is better. However from the comparison between the other numbers it is still clear that hypothesis 1 is justified to be true.

H2 earlier removal of main fences is effective. The comparison between OP and FP can illustrate that removal of the main fence leads to a drop of MAX, danger zone($\geqslant 7$) and ADen. There are slight increases in danger zone($\geqslant 5$) and ACon but the difference is relatively small. Thus we can claim that hypothesis 2 needs more evidence to be supported or rejected.

H3 utilization of the side ramp as entering and exiting channel is effective. The comparison between RP and OP illustrates that RP has a large drop in almost all the numbers compared to OP. Hypothesis 3 is true.

H4 separating the inflow and outflow by using the side ramp as exit channel is effective. The SP group outperforms in all of the simulation results, supporting hypothesis 4.

H5 applying the removal of the police cordon to the other three strategies gives an effective improvement. For group SP/SNP and group RP/RNP, the removal of police cordon improves the performance. Only for FP/FNP it is not so obvious that the removal of police cordons is effective. The simulation result exhibits a drop in D($\geqslant 7$) but an increase in D($\geqslant 5$). The other numbers are very close. So this hypothesis needs more evidence to be supported or rejected.

From the above analysis it is clear that separation of flow with the police cordon removed is the optimal crowd management strategy for Love Parade disaster in 2010. In the following Chapter 5, we will examine if the optimal strategy can be tested and compared on individual's level.

Chapter 5

Experiment

The simulation result from Chapter 4 reveals the potential effectiveness of crowd management strategy and its influence on the reduction of casualties. Another factor in crowd disaster is the pedestrian's decision making during the movement, which is strongly linked to their stress level. Stable emotional state can help individuals to make rational decision and avoid the irrational aggressiveness caused by panic. In this experiment, an approach which can help us to examine the effectiveness of the management strategies on individual's level was designed with first person perspective VR device. This chapter aims at studying the connection between the crowd simulation and the stress level of the participants.
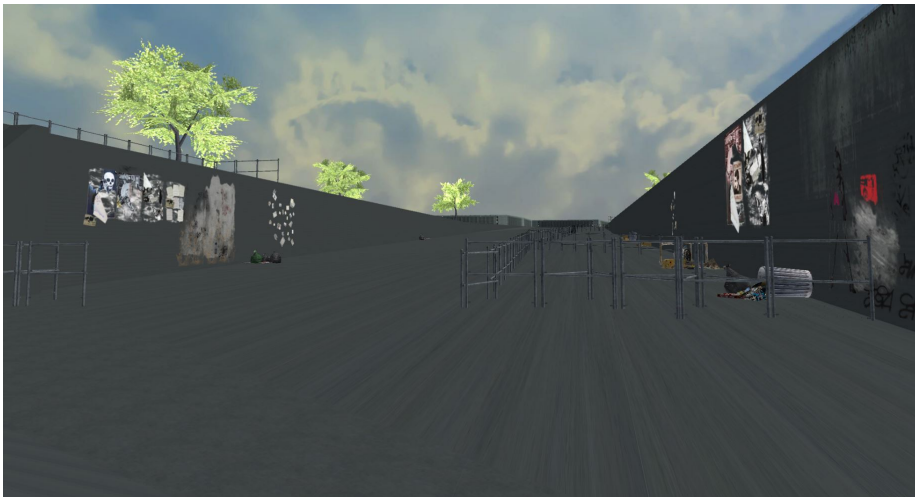
Figure 5.1: Control group replay

5.1 Experimental Design

In order to study the influence of each strategy, a comparison of stress level between individuals is made among three groups: Control group, Original group and Optimal group. Each participant was randomly chosen to belong to one of the three groups and illustrated the same replays within the group.

Independent variable The independent variable in this experiment is different management strategy, three group settings are chosen:

Control group serves the role of referential indicator for stress level. In this group the first person replay was generated with the setting of no other pedestrian(see Figure 5.1).

Original group is an authentic projection of the crowds movement in real event. Incoming flows of the crowds came from two directions of the tunnel and there was an outgoing flow that using the main ramp to exit the festival area(see Figure 5.2). The simulation trace was from previous simulation in chapter 4.

Figure 5.2: Original group replay

Optimal group is chosen based on the simulation results. Among the combination of the simulations, the strategy that removes the police cordons and separates the incoming flow and outgoing flows by using the side ramp is the best strategy that leads to smallest number of congestion and casualty. It is obvious that the utilization of the side ramp as an independent exit-only pathway can drastically improve the crowds movement efficiency. Thus we

Figure 5.3: Optimal group replay

choose to use one side of the tunnel as the incoming flow entrance point and the side ramp as the exiting point. In this case both flows of the crowds would not confront pedestrians with opposite moving direction(see Figure 5.3).

Each group has 4 total replays with pauses in between. Each replay lasted exact 2 minutes. Experimenter helped the participants to put on and off the Head-Mounted Display(HMD) between the replays. The group settings can be summarized in Table 5.1.

Table 5.1: Group comparison

Group	With companion crowds	With congestion	With confronting flow
Control	No	No	No
Original	Yes	Yes	Yes
Optimal	Yes	No	No

Dependent variable Dependent variable of this experiment is the stress level of each participant. It consists of the self rated Short Stress State Questionnaire(SSSQ) and the objective physiological measurement, in which the latter was measured by professional life science research device PowerLab[22].

Hypothesis We expect that the Optimal group with separated crowd flows has lower stress levels than the group with the Original crowd simulation but higher stress levels than the Control group. This hypothesis is drawn from real life experience, where human usually would feel more stressed

and panic when there is highly loaded crowds[27] and congestions. One of the goals of this study is to examine if this correlation can be replicated in VR environment, especially with HMD.

5.2 Method

5.2.1 Participants

Participants were master or PhD students from ETH and UZH. There were 21 participants in total, randomly assigned into three groups. There were 12 females and 9 males, aging from 23 to 34 (average age 26). Participants voluntarily joined this study and they were indicated to avoid alcohol and cigarette consumption before the experiment. In the end of each experiment, participant received a piece of chocolate as reward, which was informed before the experiment.

5.2.2 Material

Software framework

Introduction of EVE Experiments in Virtual Environment (EVE) framework is developed by Chair of Cognitive Science, ETH Zurich to fulfill the needs of researching in VR. The framework helps researcher to facilitate the integration and evaluation of their designed experiments. It helped our study in terms of scene management, data recording and statistical analysis. Its major functions (see Figure 5.4) include "simplifies the interaction with different physical and virtual sensors" and "manages data gathering during the experiment as well as access afterwards"[15]. The overlapping of the use of Unity3D, LabChart and the VR environment makes it a well suitable tool for our experiment.

Database A Mysql database was established firstly to provide data storage. All of the questionnaire answers, position information and timestamps were directly logged into the database(see Figure 5.5). The database was developed in form of supporting Unity game engine and $C^{\sharp}$ Mysql API.

Experiment management with EVE The primary procedure that needs to be managed is the game scenes and the transitions between them. There are the five types of scenes: Questionnaire, Task instruction, Baseline video, Training and Experimental replay. Each scene was adapted in need of their features during the experiment procedure. The load manager of EVE was in charge of associating diverse scenes and keeping record of timestamps. Each scene's starting time and ending time was logged into database. In addition in the experimental replay scenes, the exact start and end time of
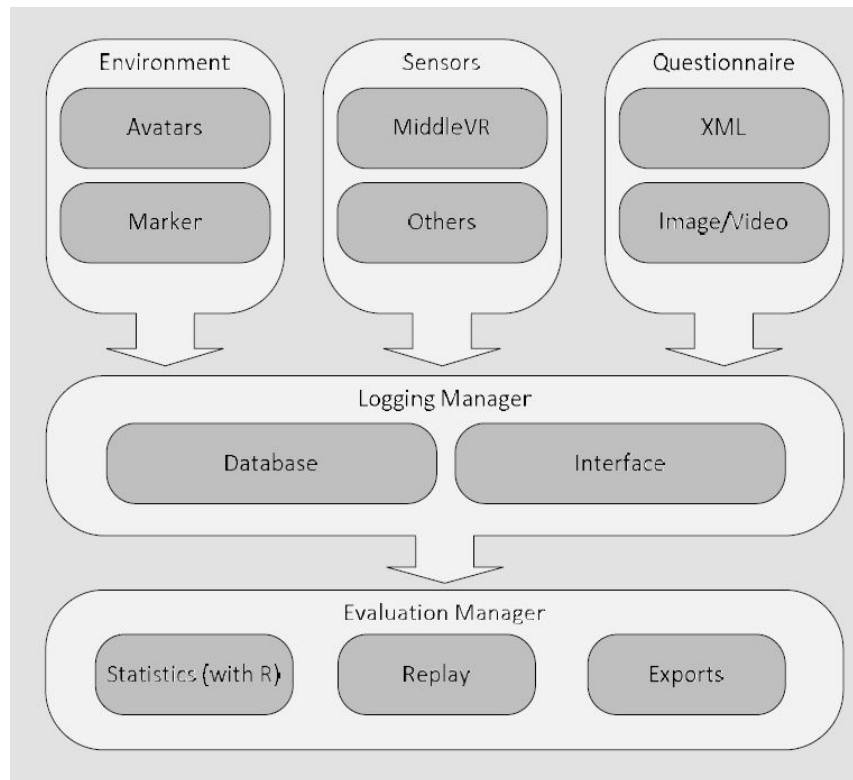
Figure 5.4: EVE framework structure

two minutes long replay were also logged, in order to match the LabChart recorder files. Plus in the game play scene the position and direction of the camera were logged as well.

Questionnaires were loaded in advance into the database and remained constant between the experiments. participant's answers were stored into the Database and could be exported at any time in the Evaluation scene, in which the position information and time record could be generated as well. Moreover after the whole experiment procedure, the timestamps in each scenes needed to be loaded into the LabChart file, in order to mark the symbolic time slots. Evaluation scene was also in charge of that function.

Other experiment settings The experiment was conducted within the Neuro lab located in ETH IFW building. Physiological data such as Skin Conductance Level (SCL) and Heart Rate Variability (HRV) was collected by Power-Lab 16/35 from ADInstruments. An upper computer software LabChart(see Figure 5.6) was in charge of the data monitoring and calibration. The sampling rate was preset to be 1000 Hz and the exported data was in form of .adchit(SCL and HRV) and .txt(SCL).
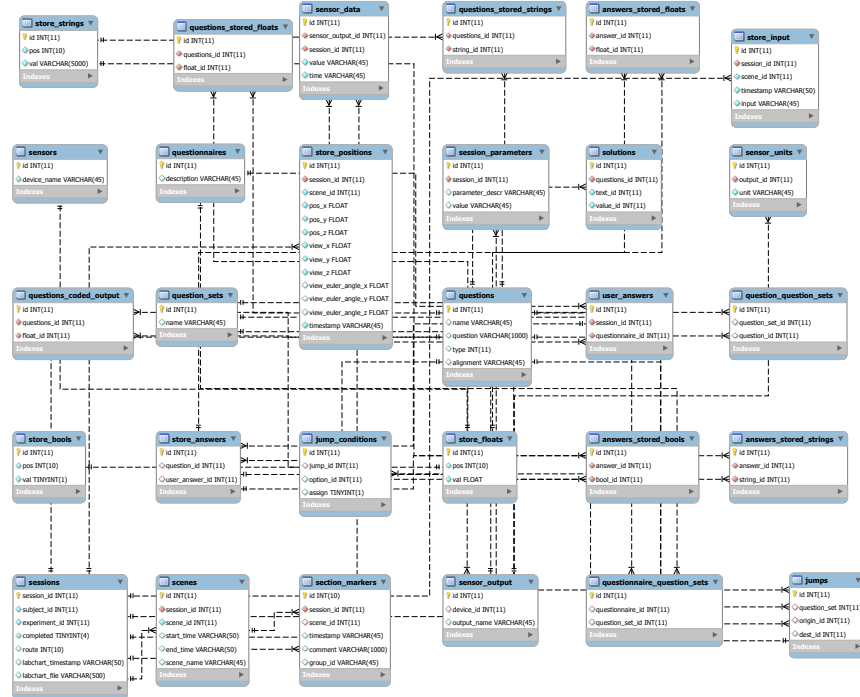
Figure 5.5: Database schema

Demographic questionnaire General demographic information was firstly collected before the experiment started. The participants were required to provide their gender, dominant hand, vision condition, contact wearing condition and VR device experience.

Video game questionnaire This questionnaire was a combination of video games experience inquiries. participants were asked about video games that they had played during the past year. They would choose between how many hours a week have they played games in each of the following categories, such as First people shooters, Action sports games, real time strategy games and VR games and so forth(See Appendix A).

SSSQ questionnaire As a sensitive indicator of states and state changes, the Short Stress State Questionnaire contains only 24 items and provide beneficial such as less response fatigue and usefulness of "investigation of stress in settings with time constraints"[13]. Helton *et al.* improved the model by reducing the influence of the long nature of the DSSQ, which contains 90
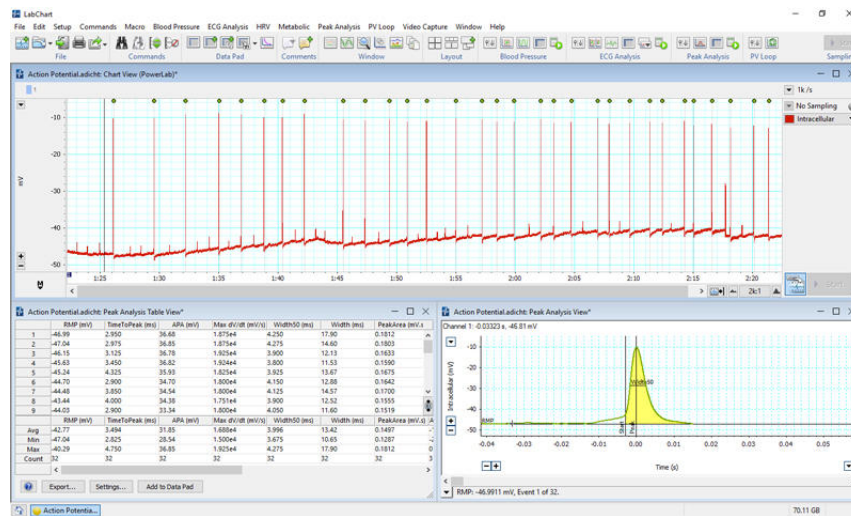
Figure 5.6: Skin conductance measurements with Powerlab
[22]

items in the latest version[25]. The questionnaire was conducted in the form of pre questionnaire and post questionnaire.

Replay mechanism   The replay of the simulation is a parallel project that shares certain similarities with the Love Parade simulation in Chapter 3. It shares the use of the 3D environment and partial crowds steering mechanism. In spite of the difference of the crowds generation and time management mechanism, the major supplement in the replay mechanism is the EVE and the Oculus integration. Some scene materials were from Unity online Asset Store. As illustrated in Figure 5.7, the tunnel was decorated and covered with textures. This could help to reduce the motion sickness of the users in HMD. Replay mechanism can be expressed with two primary modulo: **First people modulo** and **Crowds modulo**. **First people modulo** is necessary for a first person game object for it is in favor of providing a trace for Oculus HMD camera. This object was chosen among the simulated crowds and ensured to follow the exact same trace of that chosen pedestrian. **Crowds modulo** depended on the needs of the three groups and three simulations of Control, Original and Optimal when scenes are provoked in advance. The position and the timestamps from simulated agents were recorded and saved as raw data files. The replay game scene firstly reloaded all of the routes data into a list of traces using Record Loader, which serves the role of loading XML and storing the traces in forms of the *TimePointRecorder*. Each node consisted of the position data and the time data. Then in *AvatarAuthoring* class a register table was generated in order to keep track of the agent initializing statues. Register table stored all

Figure 5.7: Tunnel

the agents name and their traces. The agent's starting position and starting time were stored and recorded separately, in order to ensure the accuracy of the initialization procedure. All of the position data was stored in forms of *Vector*3 and timestamps in form of float.

Another aspect that we want to investigate is the connection between the individual's surrounding crowd density and the arousal level. For instance in the Optimal group simulation, the first half of the replay is not quite different from the Original group. Yet when the first person player reaches the center of the ramp, due to the incoming crowds from the other side of the tunnel and the exiting crowds from the festival areas, the surround density of the first people player would have a drastic increase. This phenomenon can be expressed with surround density number. To keep track of this data, the first people player had a Sphere Collider component attached on it, and *NeighbourRecord* class would use *OnTriggerEnter*() and *OnTriggerExit*() to oversee the crowds number in an area of 10 meters surrounding the first people player. The number of the neighbor crowds was recorded once every minute in simulated time scale(see Figure 5.8).

One major challenge of the replay mechanism is the time management. With regard to the exact projection of the simulated crowds statues, a mechanism needs to match the accuracy of the crowds position in the sake of timestamps. Since the simulation time is a convert of frame counts and elapsed time(see Chapter 3) instead of using real world time, we can not apply the exact same time manager from previous simulation. The new time manager needs to keep counting the elapsed real world time. The variable *Time.realtimeSinceStartup* from Unity native Time class can report the real time in seconds since the game started. *FixedUpdate*() used an accumulator to keep updating the game scene time by adding real time to it.

Once the replay scene is launched, the *AvatarAuthoring* class keeps track
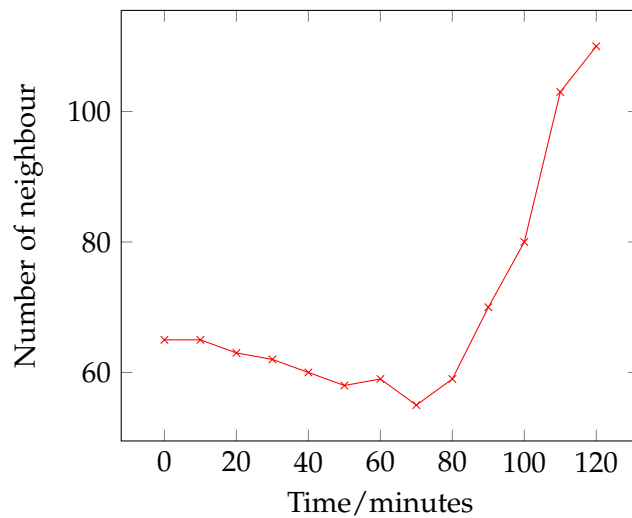
Figure 5.8: Surround density change of Original group simulation

of the real world time meanwhile initializes the agents if the system time goes beyond the agent's starting time stored in the register table. The agents were generated by cloning of the diverse avatars from an avatar pool, which consists of more than 100 distinct avatars in forms of men, women and children(see Figure 5.2). Each agent was initialized with a stack of targets pre-loaded from the simulation position data. Agent steering mechanism was taken care of by *Navmesh* agent of Unity. The agent stored a continuous targets and once they hit the current target, the next target would be dequeued from the stack and set to the Navmesh agent.

The intention of ensuring that the agent arrives at the target on the desired time requires a position calibration mechanism. This procedure was done by constantly tuning the agent's speed. The following code executed this action by calculating the speed based on the division of the distance between the target and the current position, by the difference between the required arrival time and the current system time:

```
calculatedSpeed = Vector3.Distance(agent.destination,
lastPostion) / ((aP.currentTarget.Time -
TimeManager.timeInH) * 3600f);
```

The speed was cast with a maximal and minimal value to keep the avatar movement in a realistic way. Meanwhile the avatar was animated with Animator component of the agent. Animator was synchronized with the speed and direction of agent's movement. The position calibration mechanism was able to generate the agent replay traces with respect to the desired timestamps. When the agent reached its final destination they would be destroyed to save computation memory.

Oculus This study used the newest generation Oculus Rift and a High-performance ALIENWARE AREA-51(CPU:Intel Core i7-6950X; GPU: Dual NVIDIA® GeForce® GTX™ 1080) to set up the hardware environment. The developing environment was Oculus PC SDK 1.10.0 and Unity 5.4.0f3 (64-bit). Including the Oculus into the replay scene required Unity editor level integration. *OVRCameraRig*, which controls HMD camera rendering and the head tracking, was added to the first person controller game object. *OVRManager* serves as the interface between game scene and the camera. Its *UserPositionTracking* parameter was disabled to make sure the participant's head movement would not be reflected into the game scene, ensuring the steady position of the replay trace. The replay required no input so *OVRPlayerController* was not activated. Diverse optimization strategies of Oculus were applied to the replay scene to make sure satisfactory graphic performance. For instance Anti-Aliasing was activated to smoothen the image and reduce jagged edges. Similar methods such as Occlusion Culling and light baking mentioned in Chapter 3 were deployed as well.
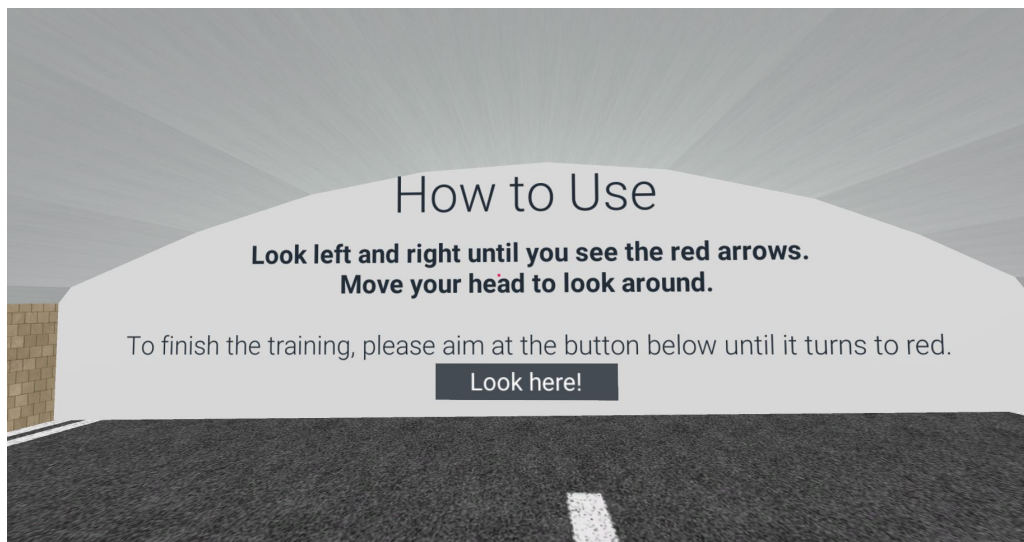
Figure 5.9: Experiment training scene

5.2.3 Procedure

Testing happened in NeuroLab located in IFW building of ETH. Before the experiment each participant was assigned randomly into one of three groups: Control group, Original group and Optimal group. After the participants arrived at the lab, they were shown firstly a description of the experiment and a consent form(See Appendix A) to sign. After given the mouse and keyboard, the participants were free to start the questionnaires.

Subsequently the experimenter helped the participants to attach physiological sensors on their fingers and chests. A short training was given after the questionnaire. The training process pledged the participants to get familiar with Oculus HMD. It also helped experimenter to be able to aide the participant to put the device on in a comfortable way. The experimenter communicated with participant to ask if they have a clear view of the camera and if they were comfortable with the Oculus. This process also made sure the experiment was conducted in a fair condition for each participant. The participants were required to look left and right firstly, then gaze at a specific button to terminate the training session(Figure 5.9). The training was set up in the same virtual environment of the replay. After the training a seven minutes long nature video was then illustrated to the participants. Its purpose was to calm the participants down and collect baseline physiological data to facilitate the between participant comparison. They were asked to remain their upper bodies and hands with sensors attached as still as possible during the whole experiment.
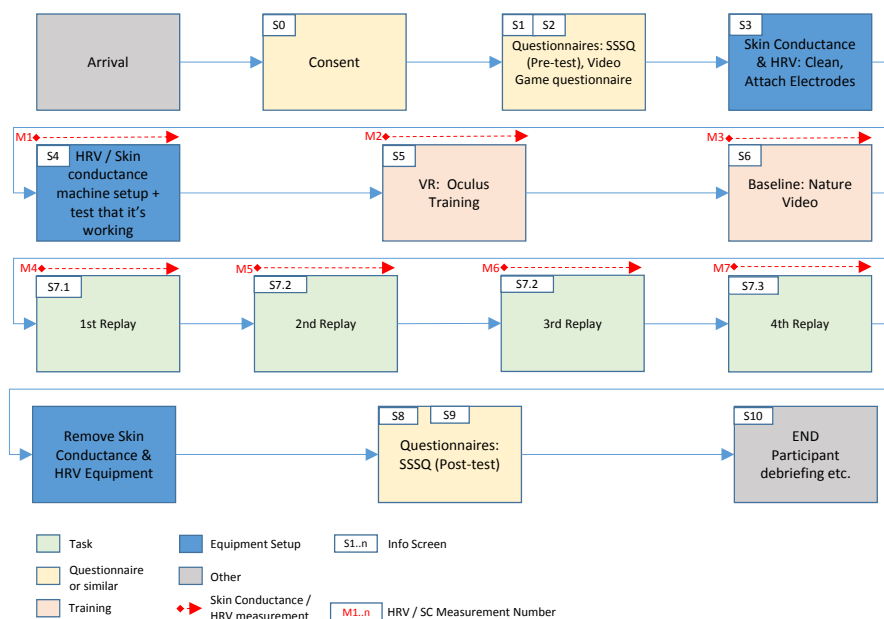
Figure 5.10: Experiment protocol

There were 4 replays in total. Before and after each replay the experimenter helped the participants to put on or off the Oculus. Each replay lasted exactly 2 minutes. Between the replays the participant has a small break. They were told that in case they felt uncomfortable or motion sick they could alert the experimenter to terminate the replay anytime. During the replay they had the freedom to move their heads and look around. After all of the four replays, a post questionnaire was provided to assess their stress state change. As represented in Figure 5.10 during the experiment the software took care of the physiological measurement, machine calibration and game scene management. The participant could keep proceeding without interruption unless there was device calibration needed. The info screens provided sufficient instruction for the participants to be aware of what stage they were at and what they should expect. Each info screen had a mandatory 10 seconds of reading time before they could continue, in order to make sure the participants would not skip any instructions by accident or on purpose.

5.2.4 Data collection and analysis

The SSSQ questionnaire needs to be processed after collection. The scores were loaded with factor table indicated in Appendix A.

$PreScore = PreQuestionnaireChoice * PreFactor$

$PostScore = PostQuestionnaireChoice * PostFactor$

The each score of Engagement, Worry and Distress are calculated in the forms as following. $factor$ include the three categories of Engagement, Distress and Worry. τ is the standard deviation of the PreScore.

$$Score(factor) = \frac{PostScore(factor) - PreScore(factor)}{\tau}$$

Factor loadings divide the importance of each item into three factor categories (See Appendix A). 24 items are categorized into three factors in order to split the analysis in a reasonable way. The Distress factor gives an indicator of how the individuals are affected in negative ways; Engagement indicates the alert and active items of the individual; Worry factor contains "self-focused attention, self-esteem, and cognitive interference items" and "appears to be entirely cognitive"[13]. Thus the 24 items are processed separately to give an independent indicators in three dimensions. This experiment focuses mainly on Engagement and Distress. The former is strongly influenced by the interests of the task, and the latter is related to objective task characteristics[25]. The Distress can also be used as a metric of how demanding the task is, which is the emotional arousal level of the participant in this study. Although Worry is less sensitive than Engagement and Distress, it can be used as an indicator of the reaction of the environment since

it correlates with low controllability and emotion-focus[25]. Engagement can be interpreted as how the participant react or is alerted by the environment and the crowds in this study, while Distress and Worry illustrate how the participant is influenced by the environment and crowds. High positive change of Distress and Worry can be used as evidence of anxiety and tensity.

Another aspect of data processing is the EDA data. Once raw EDA data was collected with PowerLab it would then be exported in the form of .txt files. Third party open source software provides raw data analyzing solution directly on the exported files. Using softwares such as Ledalab (www.ledalab.de) for the analysis of skin conductance data can simplify the data process procedure. Data extracting from the continuous phasic activity can be applied using Ledalab. Benedek et al.[1] studied that the phasic activities required a deconvolution of the skin conductance data. The convolution of the process is described as convolution of the driver with IRF, which is the course of the Impulse Response Overtime. Tonic skin conductance activity such includes Skin conductance level and Phasic skin conductance activities include phasic activity such like Skin Conductance Response (SCR). Those two together consist of the skin conductance raw data. The convolution of the drive with IRF can be presented as[1]:

$$SC = SCTonic + SCPhasic = (DriverTonic + DriverPhasic) * IRF$$

furthermore it can expressed as,

$$SC/IRF = DriverTonic + DriverPhasic$$

Hence once we have the tonic driver, the phasic driver can be calculated by subtraction from the initial value. Moreover optimization is done with gradient descent method[36] to achieve criterion of improvement. This process would end up with providing the following essential results from each phase(e.g. example training, baseline and replays) of the EDA data.

Ledalab gives the following decomposition results directly[23]:

CDA.nSCR Number of significant (= above-threshold) SCRs within response window (wrw)

CDA.Latency Response latency of first significant SCR wrw [s]

CDA.AmpSum Sum of SCR-amplitudes of significant SCRs wrw (deconvolved from corresponding phasic driver-peaks) [muS]

CDA.SCR Average phasic driver wrw. This score represents phasic activity wrw most accurately, but does not fall back on classic SCR amplitudes [muS]

The data analysis focuses on two primary parameters: nSCR and AmpSum, since they represent the phasic indicators of the SCL.

Baseline and Law of Initial Value(LIV)   Raw EDA data contains individual differences and momentary level of physiological functioning. To eliminate this effect, a comparison between the experimental data and the baseline of participants is necessary. In this study a baseline of 7 minutes of nature video playing was collected before the experimental replay started(Chapter 5.2). The characteristic of the video ensured that the physiological data collection during this stage could be used as a baseline. However the utilization of the baseline must follow a law defined by Wilder[38]. He described the law as the connection between the baseline and the after experiment degree of response. The high level of response in baseline implies the smaller response to the raising stimulation but larger response to depressing stimulation[38]. Benjamin's further research revealed that law of initial value can be stated as the correlation between pre-stimulus level and the response to stimulus, in which negative correlation implies participants with high initial level present minor response to stimulus but participants with low initial level presents major response to stimulus[2].

In order to apply Benjamin's method of statistical treatment, a best fit linear regression model needs to be drawn in order to extract the pattern of the data. A Matlab function of $fitlm(x,d)$ can be utilized to easily generate the needed model, in which x represents the baseline level and d represents difference between the response level and the baseline level. The baseline level x is normalized with respect to time. An estimated regression coefficient called bDX will be generated with the definition of $bDX = \sigma D/\sigma X$. After being applied with the linear regression to the initial data, the result data would follow the Law of Initial Value if the estimated coefficients of the linear model is negative. The estimated coefficients value could be used again to eliminate the baseline from the post stimulus response. Each participant's final score is defined with a covariance model as

$$Score = (d - bDX * x) \tag{5.1}$$

This score would have zero correlation with the baseline level, which fulfills our original purpose of the utilization of the baseline[2]. The covariance model would be applied for both nSCR and AmpSum score calculation.

Chapter 6

Results

The experiment results could help us to decompose the relationship between the management strategy and the participant's stress level. Data processing and analyzing methods are provided in previous Chapter 5. Participant's data was categorized into three groups and processed separately. No motion sickness or other uncomfortable syndromes were reported during the experiments by participants, which demonstrated the usage of decorative texture and short replay setting.

6.1 Experimental results

6.1.1 Short Stress State Questionnaire

SSSQ provides the subjective evaluation from the participants themselves. The results are illustrated in Figure 6.1. The error bar in the graphs is the standard error of the mean. It reveals the following facts:

1. The Control group has the highest Engagement effect (Mean=0.51, SE=1.65), whereas the Original group has lower Engagement(Mean=0.31, SE=1.18) effect. The Optimal group has negative Engagement effect (Mean=-0.745, SE=0.68).

2. Original group stands out in both Distress(Mean=0.55, SE=1.39) and Worry(Mean=1.26, SE=0.89) categories. While Optimal group worried less after the experiment, and had a slight increase in Distress level. Control group had the effect of making participants distress less and worry less.

3. A one way ANOVA was used to justify the results. The F value(Engagement F=0.29, Distress F=0.15, Worry F=0.74) is lower than Critical F value. High P value(Engagement P=0.74, Distress P=0.85, Worry P=0.48) implies that this experimental statistic is insignificant. Due to the rea-
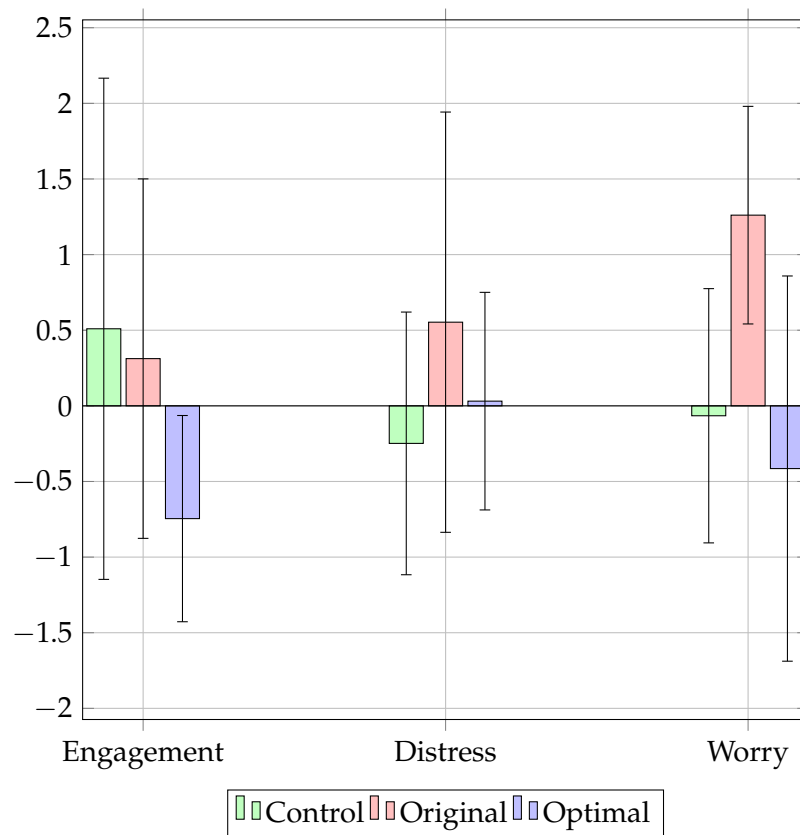
Figure 6.1: SSSQ questionnaire results

son of small size of the sample(See Chapter 5), this is comprehensible. However the results can still be utilized to analyze the trends.

This results provide some evidence that Original group has higher stress level change in terms of Distress and Worry. The Optimal group provides the context that leads to the worst Engagement but it helps to reduce the Worry and Distress level, comparing to the Original group. Control group does not differ largely from the Original group in terms of Engagement, but makes participants distress and worry less.

6.1.2 Electrodermal Activity

The objective measurement of the physiological data of the participants are processed in methods presented in Chapter 5. nSCR data depicts the frequency of the skin conductance response activity within the time scale of two minutes. It can be seen in Figure 6.2, Original group has the highest number of nSCR (Mean=31.69,SE=4.12). Despite that the mean values of all of three groups are close, the general trends can be notified with comparison
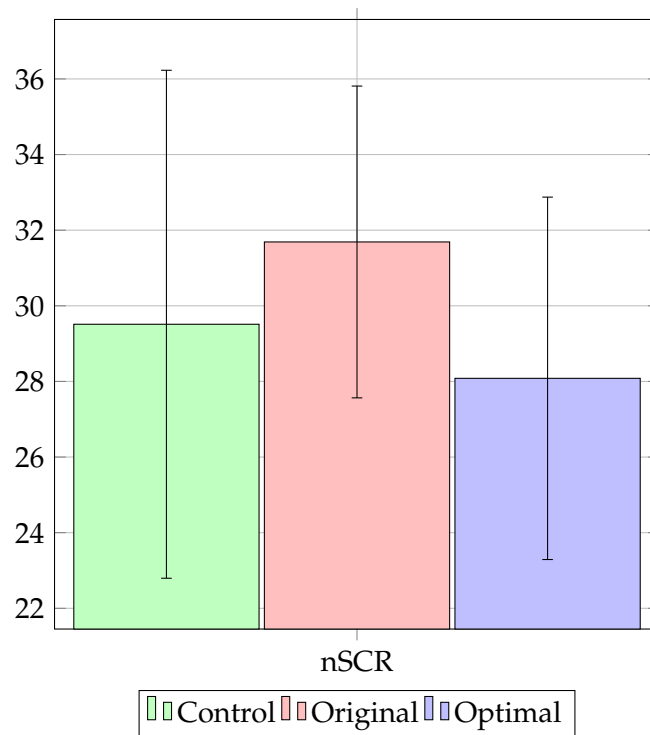
Figure 6.2: Number of significant SCRs

between Original and Optimal group(difference=3.6).

Sums of the SCR amplitudes of significant SCRs(AmpSum) are illustrated in Figure 6.3. In this indicator, the Control group has almost doubled value compared with the Original group. Whereas the Original group still has higher sum number than Optimal group(difference= 2.38).

A one way ANOVA was applied to the above data as well and similarly it reveals an insignificant statistic result. The nSCR(F=0.12, P= 0.89) and AmpSum (F=1.9, P= 0.18) both are not significant. It can also be explained by small sample size.

Two box plot in Figure 6.4 and Figure 6.5 were made below to illustrated the standard deviation of the statistics. In both nSCR and AmpSum, the Control group has the highest standard deviation. The data from Control group is not well distributed and implies heavy diversity within the group. The median of nSCR reveals different result from the mean value but the general box range matches with the mean. Nevertheless this result reflects that Original group tends to have higher emotion arousal level than the Optimal group. The Control group surpasses the Original group in AmpSum but falls behind the Original group in nSCR. The Optimal group has the lowest mean value in both nSCR and AmpSum.
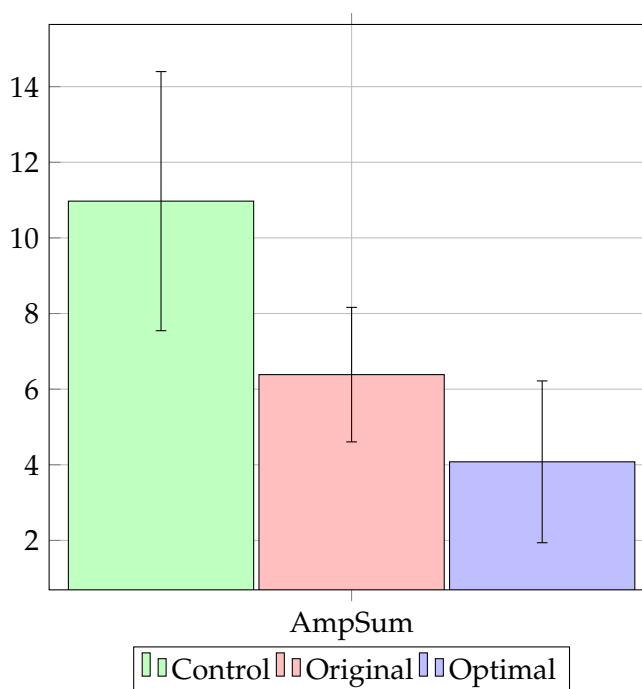
Figure 6.3: Sum of SCR-amplitudes of significant SCRs

6.2 Other correlation and findings

A cross correlation has been applied to the following parameters: (See Table 6.1)

Room humidity(Room), Gender, Age, Shortsighted or Farsighted(Vision), VR device experience(VR), SSSQ(Engage, Distress and Worry), number of significant SCR (nSCR) and sum of amplitudes(Amps).

Table 6.1: Cross correlation

	Room	Gender	Age	Vision	VR	Engage	Distress	Worry	nScr	Amps
Room	1									
Gender	0.22	1								
Age	-0.118	0.127	1							
Vision	-0.19	0	-0.14	1						
VR	0.499	0.146	-0.074	-0.044	1					
Engage	-0.08	-0.261***	-0.091	-0.174	-0.17	1				
Distress	-0.085	0.166	0.331**	-0.098	-0.386**	-0.414**	1			
Worry	-0.068	-0.293**	-0.072	-0.364**	-0.193	0.607*	-0.181	1		
nScr	0.081	-0.178	-0.230***	-0.085	-0.157	0.185	0.099	-0.084	1	
Amps	-0.407**	-0.104	-0.270	-0.140	-0.338**	0.222***	-0.156	0.162	0.02	1

* $p < 0.05$, ** $p < 0.2$, *** $p < 0.5$

Due to the nature of the small sample, the alpha for p value is not defined thus the significance level of the statistics are only an auxiliary consult-
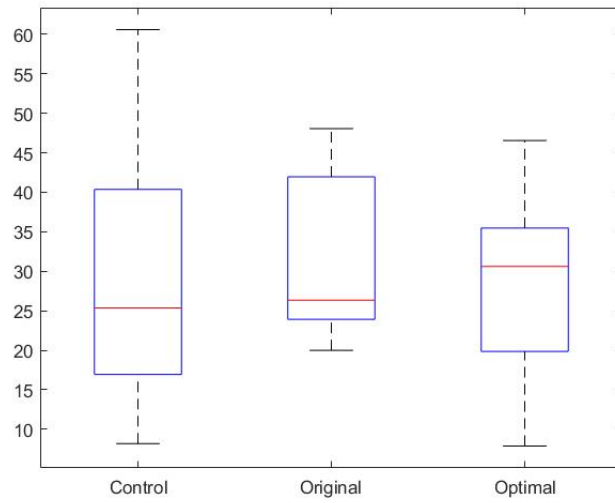
Figure 6.4: Number of significant SCRs
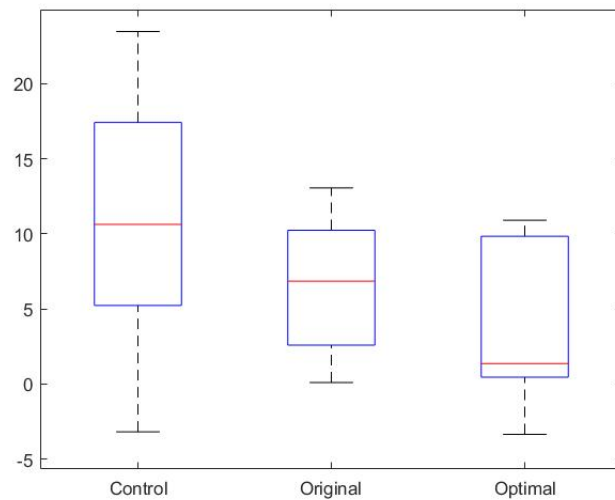
Figure 6.5: Sum of SCR-amplitudes of significant SCRs

ing. However the obvious correlations, for instance the correlation between Room humidity and Amps, the correlation between VR experience and distress, can not be neglected.

participants demographics One of the findings is the correlation between room humidity and sum of SCR amplitudes. The higher the room humidity is, the lower sum of the AmpSum is. Vision condition seems to have negative

influence on Worry. Participants who does not have vision corrections tend to worry less. Gender has a correlation with SSSQ. Females tend to has lower Engagement and Worry state change than males. Older participants have higher distress level and skin conductance level (both nSCR and AmpSum). Participants with previous VR experience tend to have lower stress level.

Stress level measurement   The only correlations that have been found between the SSSQ and skin conductance level are the correlations between Engagement and the other four factors such as Distress, Worry, nSCR and AmpSum. Higher Engagement change implies higher Worry, nSCR and AmpSum level, but it also tends to correlate Distress in negative way.

Chapter 7

Discussion

7.1 Results analysis

Short Stress State Questionnaire   There are two possible interpretations from the SSSQ results. First of all, conclusion can be drawn that the Optimal strategy can reduce the stress level comparing to the Original strategy group. Both Worry and Distress metrics indicate that participants felt more nervous and worried in the Original simulation where the crowds entered the field from two directions, facing the exiting crowds at the same time. The Optimal group even has the effect of relieving the stress, considering that the Worry measurement has an negative change. This comparison gives the evidence that when participants experience a crowded scenario in VR environment, their stress level is influenced in a similar way of real life. In Optimal group the participants are only shown with the accompany crowds that moving in the same direction with them without facing opposite direction crowds. This arrangement seem to help them to feel safe and unworried. On the other hand the Engagement metric gives another perspective about this task. It is clear that the Engagement level is way higher in Original group, which implies that participants are actually more activated in Original group than the Optimal group. The Control group gives us an interesting result to be interpreted. This group has the highest Engagement among the three groups. Although in Distress metric it matches our hypothesis, Control group tends to have higher Worry level change than the Optimal group. This gives the trends that the congestion actually cause more engaging effect than stressful. It leaves us the room to potentially investigate whether congestion condition could really motivate participants to be more actively involved. In general if we focus on the Distress and Worry level mostly, the Control group serves its role in this two scales.

Skin Conductance data   The number of significant skin conductance response (nSCR) emphasizes the influence of the Original group. High nSCR number of the Original group clearly indicates that participants have more often emotional arousals during the Original replay. The replay of highly crowded movement and formation of congestions could possibly be the cause of it. The sum of amplitudes of significant SCR (AmpSum) also support this conclusion. The Original group has 59.5% higher AmpSum than the Optimal group. The AmpSum metric implies the level of the emotional arousal, which in this case can be interpreted as how severe is the participants emotional reaction to the environment. Combining with the crowds arrangement and the congestion level described in Chapter 4, this EDA data also matches with our hypothesis.

Once again the Control group reveals different result from the hypothesis. The Control group even has the highest AmpSum value among all of the three groups and the nSCR is higher than the Optimal group. The box plot(Figure 6.4 Figure 6.5) illustrates another explanation of this effect. The deviations are very large in Control group, different from the other groups. This could also be used as explanation of the unexpected effect from the Control group. The samples from the Control group is simply too sparse to be meaningful. Combined with SSSQ result this occasion can be explained with the peculiar setting of the Control group. Although from SSSQ result we can tell that the Control group did not serve the role of making participant more nervous, it leads to more frequent and higher emotional arousals. There are also several other possible explanations: **1.** in the Control group, there was no other crowds to block the view of the participant so they actually had broader field of view and better visibility. The environments are identical in all of the three groups and it contains a lot of details such as garbages on the ground and trees. The participant can be attracted by these details, which would cause emotional arousal; **2.** in the control group the participant walks along by themself. This may cause them to feel lonely or begin to wonder what would happen next. This mental activity could be reflected with emotional arousal as well since it involves excitement and nervousness; **3.** the high Engagement level from SSSQ also supports the point that Control group can make people have the feeling of finishing this task by themselves actively. While in the other two groups, being surrounded by other crowds would make participant feel more like being an observer than being a pedestrian. In general this is an interesting finding and can be improved by introducing another control group with different setting in future studies.

Hypothesis   In conclusion, the first part of the hypothesis, *the Optimal group with separated crowd flows has lower stress levels than the group with the Original crowd simulation*, is true. However the second part of the hypothesis, *the Op-*

timal group has higher stress levels than the Control group, needs more evidence to be supported or rejected.

Other correlations   From the correlation results it can be conducted that Worry and Distress level are not correlated with the physiological measurement. This is expected since the skin conductance level is only a reflection of the emotional arousal instead of an indicator of one specific kind of emotion. The correlation between Engagement level and AmpSum represents these two indicators can be highly linked. The correlation between Engagement, Worry and Distress should not be interpreted too much since these three categories are designed to be independent variables for stress state analysis.

The correlation between room humidity and AmpSum is comprehensive due to the nature of the skin conductance measurement method. It includes two fully exposed to the air medal sensors and the room humidity could have an influence on its performance. The gender influence is another interesting effect. Females tend to be more worried and engaged than males. Considering that the males have more game experiences, this effect can be possibly caused by the female's fresh new experience to VR. The participant's vision conditions influences the Worry level. Due to the fact that the HMD requires that the user to be able to focus their eyes in 3D virtual environment, participants with worse vision ability could have higher worry level. Previous experience with VR also can help to reduce the Distress level and AmpSum level. For most of the participants the first time to put on HMD device will certainly have an influence on their stress level and performance. Elder participants' high levels of Distress and skin conductance levels also reveal the fact that HMD is easier to be adapted by younger users. The elder participants may have less experience with games and may tend to avoid exciting activities comparing to younger participants. The experiment arrangement has a tendency to be a scenario of gamification so this fact is also comprehensible.

7.2 Discussions

The overall result between the Optimal group and Original group is an exact match with the hypothesis. The Control group does not serve the role as a baseline standard due to several reasons analyzed previously. This finding also helps us to reconsider about introducing another Control group, which can be done for the next study. The new control group can be rebuild based on various number of crowds(e.g. a group with only one tenth or one fourth of the crowds of the Original or Optimal group). Another idea for a new control group is to eliminate the influence of the environment, which can

be done by rebuilding an empty tunnel and removing the garbages on the ground.

The finding of this study is helpful for crowd simulation that uses VR or HMD technologies. It implies that better crowd management can reduce participant's emotional arousal level in VR environment. In future crowd simulations, especially for crowd disaster simulations, researchers can use the first person perspective to replay the disaster in order to examine the feasibility of the strategy. Different crowd management strategy can be tested on real participant to inspect whether it would lead to panic, which is usually a possible cause for crowd disaster. With this research pattern of Crowd management ⇌ Crowd simulation ⇌ VR replay, decision makers can apply their environmental settings in advance to check if they are indeed secure and reliable. In this specific study, Helbing and Mukerji's proposed strategies [12] have been illustrated to indeed function as methods to avoid the disaster in Love Parade of 2010. The removal of the police cordons and the separation of inflow and outflow greatly reduced the casualty. It also demonstrated that the combination of optimization of strategy could be tested successfully with crowd simulation tool. *SocialForceModel* has presented its usability in high number level of crowd steering.

The physiological data gives us reliable support for emotional arousal level measurement. Given the special hardware setting of HMD, the user experience of Oculus is often a challenge for researchers. Even though there is no significant correlation between the self rated stress level and the emotional arousal level, the between group comparison gives us an evidence for evaluating crowd management strategy. In particular it illustrates that worse crowd management strategy leads to higher emotional arousal, which can be interpreted as excitement or nervousness. The example use of skin conductance measurement can also help our future research to study user experience of HMD or VR technology.

Chapter 8

Conclusion and future works

8.1 Conclusion

This thesis focused on the crowd simulation and rationales to test crowd management viability. The simulation framework demonstrated the hypothesis and experiment results met with the original design expectations. By using physiological analysis, this study gives a support for its provided optimal management strategy. According to the observed results made during a number of experiments and the simulations, the following conclusions can be made.

The most important finding is the demonstration of the crowd simulation and management strategies. Firstly, the application of Social Force Model limits the crowds density during replayed to be presented in a realistic way that would not exceed theoretical limit. The introduction of the concept of danger zone allows us to keep track of the casualty number during the simulation, which is close to the real casualty number of the real event. This depicts that the framework has successfully served the role as a sophisticated crowd simulation tool. Secondly the management strategies have been enumerated and combined to explore the possible solutions to avoid this disaster. The optimal strategy has avoided the disaster completely by separating the inflow and outflow. The results demonstrated that organizational changes to the simulation are essential for crowd disaster prevention. The cascaded crowds navigation system of Unity Navmesh and *SocialForceModel* can be expanded to future research concerning crowd simulation.

In the second part of this study an empirical experiment was designed to investigate the management strategies' influence in an individual's level. A state of the art VR device is utilized to create an immersive first person perspective replay of the event. Combined with the simulation, three groups of Control, Original and Optimal were created to examine the correlation

between crowd strategy and the individual panic level, which is usually a factor in crowd disaster. Both subjective questionnaire and objective physiological analysis provide a reliable findings in between group comparison. Optimal strategy indeed can reduce individual stress and emotional arousal level comparing to original strategy. This finding provides a second supportive evidence for the effectiveness of flow separation strategy. Even though the error levels and confidence levels were not satisfactory, considering the fact that the data has been collected in small scale, we are confident that enlarging the data sample scale can address this.

As depicted, the methods utilized in this study provided a complete rationale for crowd disaster studies, especially for management strategy verification. The up to date Virtual Reality technologies allows us to apply instantly the simulation results to a first person experience by recreating the disaster in the experiment, which is usually difficult and unethical in real field experiments.

Certainly there is room to improve in this study. The following limitations can be avoided if redo differently.

1. At the time when the experiment was conducted there was another study also used the lab for testing. Due to the lab resource conflict and limitation of time, the number of participants was not ideal. This can explain the relatively large errors in the statistics. In the formal experiment in the future there will be 60 participants with 20 per group.

2. When there is a bottleneck during the crowd simulation, congestions were formed easily. *Social Force Model* allows the trade off of high density and low congestion, but this study requires both features at the same time. Changing to another steering model (e.g. Modified Social Force Model [19]) or finding a better way to tune the parameters may help to avoid this phenomenon.

3. The experiment part was a replay of one of the simulated traces. To create an immersive feeling of evacuation and crowdedness, an actively controllable first person game object is preferable. If we allow the participants to be able to move freely and finish the task of navigating to the final target by themselves spontaneously, the participant would feel more immersive and involved. A real time interactive Social Force Model would accomplish this feature and provide another layer of authenticity during the experiment.

4. The Control group in the experiment does not match the hypothesis of providing a baseline for the stress or emotional arousal level. This could possibly be caused by the enriched environment settings. A second control group without environmental decorations can be considered to include in the formal experiment.

8.2 Future works

There are other aspects to be explored in future works. First of all in this simulation high dense crowds were simulated in a constant way. In real crowd disaster where casualty was usually caused by stepping or suffocations, the high density area would change the crowds behavior. An interesting domain to be explored is how to scientifically convert the density number into the casualty. A casualty determination mechanism could help the researchers to establish a disaster severity estimator. Secondly, from this crowd simulation platform we could build a real time crowd forecasts system. The forecast system can help to monitor the real time crowds behaviors and then conduct predictions based on the current data. Big data analysis and machine learning techniques would help to achieve this goal. Collaborative filtering algorithms can generate the crowds flow choices based on historical data. In public spaces and large events, managers can use the alert system to prevent potential crowd disasters. When evacuation happens, temporary barriers and evacuation channels can be established in a more reasonable way, based on the current crowd situation. Last but not least, this forecasting tool can then help the organizer to improve their current environment settings. A crowd behavior analysis tool which can predict people's choice of path will benefit the decision makers in a lot of ways. For example in theme parks, the infrastructures and amusement attractions can be reallocated in a scientific way to better lead the flows. This approach would also help the visitors of the theme park to enjoy their stay by saving time. Moreover in similar conditions such as Virtual Reality experiences, because of the flexible adaptation ability of the virtual world, responsive environment alternation can happen in real time to improve the visitor's experience. For VR museum visit, we could adapt the visiting orders and room settings based on the visitor's behaviors during their visit. This kind of tool will fully extend the possibilities of VR.

Appendix

A.1 Demographic questionnaire

1. Sex? Male Female

2. Please enter your age in years:

3. Are you right handed or left handed? Right handed Left handed

4. What is your vision condition? Shortsighted Farsighted Normal

5. Are you wearing contacts now? Yes No

6. Have you played with any Virtual Reality devices before? Oculus Rift, HTC Vive, Both, I have never tried either VR device.

A.2 Video game questionnaire

Over the last 12 months, how many hours a week have you played games in each of the following categories?

Example: If you play 1.5 hrs/week, mark "1+ to 3"

1. FIRST PERSON SHOOTERS (Halo, Call of Duty, Gears of War, GTA, Half-Life, Unreal etc)

2. ACTION/ACTION-SPORTS GAMES (God of War, Mario Kart, Burnout, Madden, FIFA, etc)

3. REAL TIME STRATEGY GAMES (Warcraft, Starcraft, Command and Conquer, Age of Empires, etc)

4. TURN-BASED STRATEGY/PUZZLE GAMES (Civilization, Sims, Puzzle Quest, Bejewled, Solitaire, etc)

5. ROLE-PLAYING/FANTASY GAMES (World of Warcraft, Final Fantasy, Fable, Oblivion, etc)

VIRTUAL REALITY GAMES (Using devices like Oculus Rift, HTC Vive or PSVR, etc)

A.3 Short Stress State Questionnaire

The following questionnaire structure and factor loadings are from [14].

A.3.1 State Pre-Questionnaire

Please indicate how well each word describes how you feel At The Moment.

Not at all = 1 A little bit = 2 Somewhat = 3 Very much = 4 Extremely = 5

1. Dissatisfied 1 2 3 4 5
2. Alert 1 2 3 4 5
3. Depressed 1 2 3 4 5
4. Sad 1 2 3 4 5
5. Active 1 2 3 4 5
6. Impatient 1 2 3 4 5
7. Annoyed 1 2 3 4 5
8. Angry 1 2 3 4 5
9. Irritated 1 2 3 4 5
10. Grouchy 1 2 3 4 5

Please indicate how true each statement is of your thoughts During The Past Ten Minutes.

Not at all = 1 A little bit = 2 Somewhat = 3 Very much = 4 Extremely = 5

11. I am committed to attaining my performance goals 1 2 3 4 5
12. I want to succeed on the task 1 2 3 4 5
13. I am motivated to do the task 1 2 3 4 5
14. I'm trying to figure myself out 1 2 3 4 5
15. I'm reflecting about myself. 1 2 3 4 5
16. I'm daydreaming about myself. 1 2 3 4 5
17. I feel confident about my abilities. 1 2 3 4 5
18. I feel self-conscious. 1 2 3 4 5
19. I am worried about what other people think of me. 1 2 3 4 5
20. I feel concerned about the impression I am making. 1 2 3 4 5
21. I expect to perform proficiently on this task. 1 2 3 4 5
22. Generally, I feel in control of things. 1 2 3 4 5
23. I thought about how others have done on this task. 1 2 3 4 5
24. I thought about how I would feel if I were told how I performed. 1 2 3 4 5

A.3.2 State Post-Questionnaire

Please indicate how well each word describes how you felt During The Task.

Not at all = 1 A little bit = 2 Somewhat = 3 Very much = 4 Extremely = 5

1. Dissatisfied 1 2 3 4 5
2. Alert 1 2 3 4 5
3. Depressed 1 2 3 4 5
4. Sad 1 2 3 4 5
5. Active 1 2 3 4 5
6. Impatient 1 2 3 4 5
7. Annoyed 1 2 3 4 5
8. Angry 1 2 3 4 5
9. Irritated 1 2 3 4 5
10. Grouchy 1 2 3 4 5

Please indicate how true each statement was of your thoughts While Performing The Task.

Not at all = 1 A little bit = 2 Somewhat = 3 Very much = 4 Extremely = 5

11. I was committed to attaining my performance goals 1 2 3 4 5
12. I wanted to succeed on the task 1 2 3 4 5
13. I was motivated to do the task 1 2 3 4 5
14. I tried to figure myself out. 1 2 3 4 5
15. I reflected about myself. 1 2 3 4 5
16. I daydreamed about myself. 1 2 3 4 5
17. I felt confident about my abilities. 1 2 3 4 5
18. I felt self-conscious. 1 2 3 4 5
19. I was worried about what other people think of me. 1 2 3 4 5
20. I felt concerned about the impression I was making. 1 2 3 4 5
21. I performed proficiently on this task. 1 2 3 4 5
22. Generally, I felt in control of things. 1 2 3 4 5
23. I thought about how others have done on this task. 1 2 3 4 5
24. I thought about how I would feel if I were told how I performed. 1 2 3 4 5

A.3.3 Factor loadings

Table A.1: Factor Loading

	Pre-Distress	Pre-Engage	Pre-Worry	Post-Distress	Post-Engage	Post-Worry
Item 2	-0.03	0.46	-0.06	0.01	0.59	-0.07
Item 5	-0.05	0.36	-0.02	-0.03	0.46	0.03
Item 11	0.04	0.70	-0.06	0.03	0.76	-0.01
Item 12	0.06	0.70	-0.02	0.15	0.65	-0.06
Item 13	0.06	0.74	-0.07	0.08	0.75	-0.04
Item 17	-0.16	0.54	0.02	-0.28	0.53	0.08
Item 21	-0.15	0.58	0.09	-0.33	0.55	0.09
Item 22	-0.13	0.45	-0.02	-0.35	0.50	0.10
Item 1	0.62	-0.09	-0.03	0.70	-0.09	-0.04
Item 3	0.54	-0.16	0.13	0.59	-0.08	0.17
Item 4	0.60	-0.14	0.08	0.55	-0.10	0.19
Item 6	0.52	0.12	-0.03	0.53	0.12	0.04
Item 7	0.74	-0.01	0.04	0.83	0.08	-0.02
Item 8	0.75	-0.04	-0.03	0.76	-0.04	0.07
Item 9	0.85	-0.01	-0.05	0.88	0.01	-0.03
Item 10	0.56	-0.11	0.08	0.65	-0.09	0.09
Item 14	-0.07	-0.15	0.67	-0.07	0.04	0.62
Item 15	-0.06	-0.12	0.64	-0.10	-0.02	0.72
Item 16	-0.04	-0.12	0.54	0.04	-0.09	0.54
Item 18	0.15	0.06	0.64	0.13	-0.05	0.65
Item 19	0.10	0.03	0.71	0.05	-0.08	0.75
Item 20	0.02	0.09	0.73	0.00	-0.06	0.77
Item 23	0.10	0.24	0.32	0.08	0.06	0.57
Item 24	0.17	0.33	0.39	0.12	0.16	0.48

A.3.4 Consent Form

Please read this form carefully. Please ask the investigator or the contact people if you have any questions.

Study title: Crowds simulation in Virtual Reality

Study location: The Neurolab, IFW B

Principal Investigator's Name and First Name: Prof. Dr. Christoph Hoelscher

Participant's Name and First Name:

Participant:
I participate in this study on a voluntary basis and can withdraw from the study at any time without giving reasons and without any negative consequences.
I have been informed orally and in writing about the aims and the procedures of the study, the advantages and disadvantages as well as potential risks.
I have read the written information for the volunteers. My questions related to the study participation have been answered satisfactorily. I have been given a copy of the information for the volunteers and the consent form. I was given sufficient time to make a decision about participating in the study.
With my signature I certify that I fulfill the requirements for the study participation mentioned in the information for the volunteers.
I have been informed that possible damages to my health which are directly related to the study and are demonstrably the fault of ETH Zurich, are covered by the general liability insurance of ETH Zurich (insurance policy no. 100.001 of the Swiss Mobiliar insurance company). However, beyond the before mentioned, my health- and/or accident insurance (e.g. for the way to or back from the study location) will apply. I agree that the responsible investigators and/or the members of the ethical committee have access to the original data under strict confidentiality.
I am aware that during the study I have to comply with the requirements and limitations described in the information for the volunteers. In my own health interest the investigators can, without mutual consent, exclude me from the study.

Location, date . . . . Signature volunteer . . . .
Location, date . . . . Signature investigator . . . .

Bibliography

[1] Mathias Benedek and Christian Kaernbach. A continuous measure of phasic electrodermal activity. *Journal of neuroscience methods*, 190(1):80–91, 2010.

[2] Lorna S Benjamin. Statistical treatment of the law of initial values (liv) in autonomic research: a review and recommendation. *Psychosomatic Medicine*, 25(6):556–566, 1963.

[3] Thomas R Carretta and Guy A French. Combating vigilance decrements in a sustained attention task: Lack of support for the utility of a cognitive intervention secondary task. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 56, pages 1446–1450. Sage Publications, 2012.

[4] Parth Rajesh Desai, Pooja Nikhil Desai, Komal Deepak Ajmera, and Khushbu Mehta. A review paper on oculus rift-a virtual reality headset. *arXiv preprint arXiv:1408.1173*, 2014.

[5] Bernd Figner, Ryan O Murphy, et al. Using skin conductance in judgment and decision making research. *A handbook of process tracing methods for decision research*, pages 163–184, 2011.

[6] Jonathan L Freedman. *Crowding and behavior*, volume 1. WH Freedman, 1975.

[7] Peter E Hart, Nils J Nilsson, and Bertram Raphael. A formal basis for the heuristic determination of minimum cost paths. *IEEE transactions on Systems Science and Cybernetics*, 4(2):100–107, 1968.

[8] Dirk Helbing, Lubos Buzna, Anders Johansson, and Torsten Werner. Self-organized pedestrian crowd dynamics: Experiments, simulations, and design solutions. *Transportation science*, 39(1):1–24, 2005.

[9] Dirk Helbing, Illés Farkas, and Tamas Vicsek. Simulating dynamical features of escape panic. *Nature*, 407(6803):487–490, 2000.

[10] Dirk Helbing, Anders Johansson, and Habib Zein Al-Abideen. Dynamics of crowd disasters: An empirical study. *Physical review E*, 75(4):046109, 2007.

[11] Dirk Helbing and Peter Molnar. Social force model for pedestrian dynamics. *Physical review E*, 51(5):4282, 1995.

[12] Dirk Helbing and Pratik Mukerji. Crowd disasters as systemic failures: analysis of the love parade disaster. *EPJ Data Science*, 1(1):1, 2012.

[13] William S Helton. Validation of a short stress state questionnaire. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 48, pages 1238–1242. SAGE Publications, 2004.

[14] William S. Helton and Katharina Näswall. Short stress state questionnaire factor structure and state change assessment. In *European Journal of Psychological Assessment*, volume 31, page 20–30. Hogrefe Publishing, 2015.

[15] Christoph Hölscher Victor R. Schinazi Jascha Grübel, Raphael Weibel. EVE: A Framework for Experiments in Virtual Environments Poster. 2016.

[16] Anders Johansson, Dirk Helbing, and Pradyumn K Shukla. Specification of the social force pedestrian model by evolutionary adjustment to video tracking data. *Advances in complex systems*, 10(supp02):271–288, 2007.

[17] Marcelo Kallmann and Mubbasir Kapadia. Navigation meshes and real-time dynamic planning for virtual worlds. In *ACM SIGGRAPH 2014 Courses*, page 3. ACM, 2014.

[18] Max Kinateder, Paul Pauli, Mathias Müller, Jürgen Krieger, Frank Heimbecher, Inga Rönnau, Ulrich Bergerhausen, Götz Vollmann, Peter Vogt, and Andreas Mühlberger. Human behaviour in severe tunnel accidents: Effects of information and behavioural training. *Transportation research part F: traffic psychology and behaviour*, 17:20–32, 2013.

[19] Gerta Köster, Franz Treml, and Marion Gödel. Avoiding numerical pitfalls in social force models. *Physical Review E*, 87(6):063305, 2013.

[20] Barbara Krausz and Christian Bauckhage. Loveparade 2010: Automatic video analysis of a crowd disaster. *Computer Vision and Image Understanding*, 116(3):307–319, 2012.

[21] Tobias Kretz. On oscillations in the social force model. *Physica A: Statistical Mechanics and its Applications*, 438:272–285, 2015.

[22] LabChart. Powerlab, 2016.

[23] Ledalab. Ledalab documentation, 2016.

[24] Gerald Matthews, Sian E Campbell, Shona Falconer, Lucy A Joyner, Jane Huggins, Kirby Gilliland, Rebecca Grier, and Joel S Warm. Fundamental dimensions of subjective state in performance settings: task engagement, distress, and worry. *Emotion*, 2(4):315, 2002.

[25] Gerald Matthews, James Szalma, April Rose Panganiban, Catherine Neubauer, and Joel S Warm. Profiling task stress with the dundee stress state questionnaire. *Psychology of stress: New research*, pages 49–90, 2013.

[26] Ramin Mehran, Alexis Oyama, and Mubarak Shah. Abnormal crowd behavior detection using social force model. In *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*, pages 935–942. IEEE, 2009.

[27] Niamh A Merriman, Jan Ondřej, Alicia Rybicki, Eugenie Roudaia, Carol O'Sullivan, and Fiona N Newell. Crowded environments reduce spatial memory in older but not younger adults. *Psychological Research*, pages 1–22, 2016.

[28] Mehdi Moussaïd, Dirk Helbing, and Guy Theraulaz. How simple rules determine pedestrian behavior and crowd disasters. *Proceedings of the National Academy of Sciences*, 108(17):6884–6888, 2011.

[29] Mehdi Moussaïd, Mubbasir Kapadia, Tyler Thrash, Robert W Sumner, Markus Gross, Dirk Helbing, and Christoph Hölscher. Crowd behaviour during high-stress evacuations in an immersive virtual environment. *Journal of The Royal Society Interface*, 13(122):20160414, 2016.

[30] Mehdi Moussaïd, Niriaska Perozo, Simon Garnier, Dirk Helbing, and Guy Theraulaz. The walking behaviour of pedestrian social groups and its impact on crowd dynamics. *PloS one*, 5(4):e10047, 2010.

[31] Nuria Pelechano, Catherine Stocker, Jan Allbeck, and Norman Badler. Being a part of the crowd: towards validating vr crowds using presence. In *Proceedings of the 7th international joint conference on Autonomous agents and multiagent systems-Volume 1*, pages 136–142. International Foundation for Autonomous Agents and Multiagent Systems, 2008.

[32] Maria Pretorius, Steven Gwynne, and Edwin R Galea. Large crowd modelling: an analysis of the duisburg love parade disaster. *Fire and Materials*, 39(4):301–322, 2015.

[33] Christian Reuter, Alexandra Marx, and Volkmar Pipek. Social software as an infrastructure for crisis management–a case study about current practice and potential usage. In *Proceedings of the 8th International ISCRAM Conference*, pages 1–10, 2011.

[34] Shawn Singh, Mubbasir Kapadia, Petros Faloutsos, and Glenn Reinman. An open framework for developing, evaluating, and sharing steering algorithms. In *International Workshop on Motion in Games*, pages 158–169. Springer, 2009.

[35] Shawn Singh, Mubbasir Kapadia, Billy Hewlett, Glenn Reinman, and Petros Faloutsos. A modular framework for adaptive agent-based steering. In *Symposium on Interactive 3D Graphics and Games*, pages PAGE–9. ACM, 2011.

[36] Jan Snyman. *Practical mathematical optimization: an introduction to basic optimization theory and classical and new gradient-based algorithms*, volume 97. Springer Science & Business Media, 2005.

[37] Ulrich Weidmann. Transporttechnik der fussgänger. 1992.

[38] Joseph Wilder. The law of initial value in neurology and psychiatry: Facts and problems1. *The Journal of nervous and mental disease*, 125(1):73–86, 1957.

[39] Katja Wolff. Analyzing the influence of the environment and crowd parameters on the 2010 Love Parade disaster, 2015.

ETH

Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Declaration of originality

The signed declaration of originality is a component of every semester paper, Bachelor's thesis, Master's thesis and any other degree paper undertaken during the course of studies, including the respective electronic versions.

Lecturers may also require a declaration of originality for other written papers compiled for their courses.

___

I hereby confirm that I am the sole author of the written work here enclosed and that I have compiled it in my own words. Parts excepted are corrections of form and content by the supervisor.

Title of work (in block letters):

Crowd Simulation and Virtual Reality Experiments for 2010 Love Parade Disaster

Authored by (in block letters):
For papers written by groups the names of all authors are required.

Name(s):	**First name(s):**
Zhao	Hantao

With my signature I confirm that
- – I have committed none of the forms of plagiarism described in the 'Citation etiquette' information sheet.
- – I have documented all methods, data and processes truthfully.
- – I have not manipulated any data.
- – I have mentioned all persons who were significant facilitators of the work.

I am aware that the work may be screened electronically for plagiarism.

Place, date	**Signature(s)**
Zürich, 10 December 2016	*Hantao*

For papers written by groups the names of all authors are required. Their signatures collectively guarantee the entire content of the written paper.