



Data-Driven Modeling of Pedestrian Crowds

Doctoral Thesis submitted for the degree *Dr.-Ing.*

by

Anders Fredrik Johansson

Born: 26 April, 1981

Faculty of Traffic Sciences „Friedrich List“
Technische Universität Dresden

Referees Prof. Dr. rer. nat. habil. Dirk Helbing
 H. E. Dr.-Ing. Habib Z. Al-Abideen
 Prof. Dr. sc. pol. habil. Knut Haase

Submitted 20 October, 2008
Defended 3 June, 2009

Contents

1	Introduction	5
2	Video Analysis	8
2.1	Introduction	8
2.1.1	State-of-the Art	11
2.2	Microscopic Approach	13
2.3	Macroscopic Approach	17
2.3.1	Introduction	17
2.3.2	Hough Transform	17
2.3.3	Artificial Neural Network	20
2.3.4	Principal Components Analysis	26
2.3.5	Making Use of the Temporal Information in Videos	28
2.3.6	Parameter Calibration	29
2.3.7	Evaluation of Accuracy	30
3	Self-Organization Phenomena	38
3.1	Introduction	38
3.2	Lane Formation	38
3.3	Stripe Formation	39
3.4	Pedestrian Trail Formation	39
3.5	Bottlenecks	39
3.5.1	Uni-Directional Bottleneck Flows	39
3.5.2	Bi-Directional Bottleneck Flows	41
3.6	Stop-and-Go Waves	41
3.7	Crowd Turbulence	42
4	Flow-Density Diagram	47
4.1	Introduction	47
4.2	Measurement of Local Densities, Speeds and Flows	49

4.2.1	Data Evaluation Method	50
4.3	Empirical Findings	53
4.3.1	Dealing with Umbrellas	53
4.3.2	Relationships between Densities, Velocities and Flows	58
4.4	Fundamental-Diagram Model	61
5	Modeling and Simulation	67
5.1	Introduction	67
5.1.1	Agent-Based Models	68
5.1.2	Social-Force Model	69
5.1.3	Cellular Automata Models	69
5.1.4	Fluid-Dynamic Models	69
5.1.5	Queueing Models	70
5.2	Social-Force-Model Specification	70
5.2.1	Improved Specifications of the Social-Force Model	72
5.3	Empirical Data	73
5.3.1	Trajectory Data	75
5.4	Hybrid Data-Analysis Approach	75
5.4.1	Distance-Dependence of Pedestrian Interactions	79
5.4.2	Angular-Dependence of Pedestrian Interactions	81
5.4.3	Number of Interactions	81
5.4.4	Delays	84
5.5	Calibrating the Social-Force Model	86
5.5.1	Parameter Calibration	88
5.5.2	Parameter Analysis	88
5.6	Simulation Results	96
5.7	Pre-Movement-Time Modeling	102
5.7.1	Social Influence	103
5.7.2	Towards a Model of Pre-Movement Time	104
5.7.3	Results	106
6	Optimization	109
6.1	Introduction	109
6.2	Parameter Optimization	109
6.3	Flow Optimization by Boolean Grids	112
6.3.1	Boolean-Grid Representation	112

6.3.2 Results	119
7 Hajj, the Pilgrimage to Makkah	122
7.1 Introduction	122
7.2 Validation of the Video-Analysis Software	122
7.3 Light Dependence	127
7.4 Analysis of a Crowd Disaster	129
7.5 Large-Scale Evaluation of Videos	135
8 Conclusions and Outlook	144
A Glossary	147
B Parameters and Abbreviations	149
C Technical Details for Video Analysis	151
C.1 Technical Setup	151
D Publications	157
Bibliography	160

List of Figures

2.1	Video tracking, used to extract the trajectories of pedestrians from video recordings.	15
2.2	Samples from the three different datasets. Dataset #1 consists of 11 frames and 196 patterns (heads and non-heads), dataset #2 of 5 frames and 258 patterns, and dataset #3 consists of 3 frames and 240 patterns. The heads and the non-heads are representing 50% of the samples each.	19
2.3	Sample frames for different transforms.	21
2.4	The topology of our Artificial Neural Network (ANN), used to classify heads.	22
2.5	Errors of the classification by the Artificial Neural Network.	25
2.6	Left: Eigenvalues from the principal components analysis (PCA), ordered after size. Right: Sample data projected on the first and second principal axes obtained by the PCA.	27
2.7	Classifications of the data projected onto the k primary components, found by the PCA method.	33
2.8	The distance error between extrapolated positions and real positions after a forecasting time of Δt seconds.	34
2.9	Probability-density surface of head detections, overlaid onto the original frame.	35
2.10	The average head radius is found by locating the peak of the average value from the Hough transform.	36
2.11	The error as a function of the threshold, T' , for different values of the fraction-of-head parameter γ'	36
2.12	Comparison of the cumulative counts and flow rates of pedestrians according to our algorithm as compared to an estimation made by manually counting pedestrians for a short time period every 10 minutes.	37

3.1	Time headway distributions of pedestrians in a uni-directional stream before a short bottleneck (left) and after it (right).	40
3.2	Illustration of stop-and-go waves.	43
3.3	Long-term photograph of a pedestrian crowd in Mina, Kingdom of Saudi Arabia.	44
3.4	Typical trajectories and an analogy to earthquakes.	45
3.5	Analysis of the turbulent flow.	46
4.1	Three probability-density functions of the local densities and speeds are plotted for three different average (global) densities which reveals a large variation of local densities.	51
4.2	In order to illustrate the relation between global and local density, a circle of 100-meter radius is filled with N points according to a given global density.	52
4.3	Actual (real) density ϱ , as a fraction p of umbrellas.	54
4.4	To determine the influence of umbrellas on the measurements, simulations are carried out with the social-force model.	56
4.5	Empirical relationships between the local speed (and the local flow) and the local density.	57
4.6	Empirical relationships between the globally averaged speed (and flow) and the global density.	57
4.7	Fraction p of pilgrims carrying umbrellas as a function of time.	59
4.8	Average local speed (left) and flow (right) as a function of the local density ρ	59
4.9	The local flow-density relationship as compared to flow-density data of other studies.	60
4.10	Net-time headways T as a function of crowd density ϱ	62
4.11	Fundamental diagrams generated by Eq. (4.9), for different free speeds v^0	63
4.12	Mean net-time headway which is obtained by integrating the Gaussian distribution of local densities.	64
4.13	Fundamental diagrams generated by assuming a constant net-time headway \hat{T} as the key control parameter.	66
5.1	A simplified interaction force is assumed, which has one angular-dependent part and one distance-dependent part.	72
5.2	Snapshots of the video recordings, and corresponding trajectories.	74

5.3	Illustration of the hybrid approach, used for fusing empirical trajectory data with microscopic simulations.	76
5.4	Examples of empirical trajectories compared to their corresponding simulated trajectories and their resulting fitness values.	78
5.5	Empirically determined distance dependence of the interaction force between pedestrians.	80
5.6	Angular dependence of the influence of other pedestrians.	82
5.7	The fitness value as a function of parameter N	83
5.8	Probability-density function of the delay parameter T^d	85
5.9	Fitness surface, spanned by the parameters A and B	87
5.10	Fitness values for different parameter values A and B	91
5.11	Force curves plotted from each of the three datasets in Fig. 5.2.	95
5.12	Result from a simulation with the social-force model, of two pedestrians who are approaching each other at an 180-degrees angle.	96
5.13	The force fields resulting from the <i>Elliptical II</i> specification of the social-force model.	97
5.14	Illustration of lane formation.	100
5.15	Degree of coordination at a bottleneck.	101
5.16	Influence of neighbors and friends on the pre-movement time.	104
5.17	Probability that a wave will propagate in different directions around a person, given that this person is raising up.	105
5.18	Number of persons who have stood up, as a function of time.	108
6.1	The geometry of two different pedestrian intersection designs	110
6.2	Flowrates as functions of simulated time for the conventional intersection compared to the improved design with optimized parameters.	111
6.3	An illustration of the effect of the clustering rule.	115
6.4	An illustration of the effect of applying the clustering rule 10 times, with different neighborhood sizes d	115
6.5	The effect of the level of stochasticity in the clustering rule.	115
6.6	The result of a series of simulations of a $20\text{m} \times 10\text{m}$ corridor, terminated by a 0.55 m wide door in the middle of one end of the corridor.	117
6.7	Examples of obstacle grids resulting from the Boolean-grid optimization. .	120
7.1	Green box, used for manual head counting.	123
7.2	The Jamarat Bridge.	125

7.3	Example of the cumulative flows during the 24 hours of the 10th day of the Hajj in 1426H.	126
7.4	Investigation of sampling frequencies.	127
7.5	Snapshots from video recordings during day (top left) and night (top right).	128
7.6	Snapshots reflecting the density build-up during 90 minutes before the accident occurred.	132
7.7	The (cumulative) inflow to the Jamarat Bridge (on the 12th day of Dhu al-Hijjah, 1426H) shows that there is a decrease in the flow, before and during the crowd disaster, but the flow is never even close to vanishing.	133
7.8	Test of various quantities for anticipating the critical situation both in time and space.	134
7.9	Camera positions (green circles) during the Hajj in 1427H.	135
7.10	Graphical interfaces used during the Hajj in 1427H.	136
7.11	Snapshots of the graphical interfaces showing the current state of the system.	139
7.12	Snapshots of five video cameras.	140
7.13	Aggregated values when adding the cumulative flows, measured for 10 cameras together, and divided by 10 to get average cumulative flows.	141
7.14	Cumulative flow of pilgrims to and from the 1st floor of the Jamarat Bridge.	142
7.15	Cross-correlation of the inflow and outflow of pilgrims.	143
C.1	Snapshots of the calibration screens.	154
C.2	A photo of the calibration process.	155
C.3	Manual flow-measurement process.	155
C.4	Pictures from the Hajj in 1427H.	156

List of Tables

2.1	The chain of transforms carried out in order to obtain the locations of the human heads.	18
2.2	A summary of the three free parameters and their meanings.	20
5.1	A summary of the parameters used by the social-force model.	86
5.2	Resulting fitness values for different parameter-optimization runs.	89
5.3	Resulting parameter values for different parameter-optimization runs.	90
5.4	Correlation coefficients R for the resulting parameter values from parameter optimizations of the three datasets A, B and C.	93
5.5	Relative fitness decline, by replacing personal parameter with global parameters.	94
6.1	Optimal parameter values resulting from our evolutionary parameter optimization.	111
6.2	A summary of the parameters and their meanings.	118
7.1	Total number of pilgrims and corresponding relative differences between manual and automated counts.	137
B.1	A list of abbreviations.	149
B.2	A list of measurables.	149
B.3	A list of parameters.	150

Abstract

At the starting point of the work leading to this doctoral thesis, in January 2005, the work on pedestrians was almost exclusively oriented towards computer simulations and on evacuation experiments. Since then, there have been many studies on new methods for extracting empirical data of pedestrian movements (mainly based on video analysis, lasers, and infrared cameras), but most of the work is still focused on artificial setups for crowds moving through corridors and crowds passing bottlenecks. Even though these controlled experiments are important to understand crowd dynamics, there is a knowledge gap between these experiments and the understanding of the dynamics leading to and occurring during large crowd disasters, when sometimes hundreds of thousands or even millions of pedestrians are involved.

Numerous crowd disasters occur every year at large gatherings around the world. Unfortunately, the information about the (spatio-temporal) development of these events tend to be qualitative rather than quantitative.

Video recordings from the crowd disaster in Mina, Kingdom of Saudi Arabia, on the 12th of January 2006, where hundreds of pilgrims lost their lives during the annual Muslim pilgrimage to Makkah, gave the possibility to scientifically evaluate the dynamics of the crowd. With this video material, it was possible to observe and analyze the behavior of the crowd under increasing crowd density, leading to the disaster.

Based on the insights from the analysis of the crowd disaster described above, new tools and measures to detect and avoid critical crowd conditions have been proposed, and some of them have been implemented in order to reduce the likelihood of similar disasters in the future.

Further contributions of this thesis are to empirically evaluate many of the previous assumptions used for pedestrian simulations. These assumptions are:

1. A pedestrian avoids collisions by changing her or his walking speed by an acceleration which is exponentially decaying with the distance to the pedestrian or object being avoided.

2. A pedestrian reacts stronger to what happens in front of her or him, than to what happens behind the back.
3. The movement of a crowd of pedestrians always follows a smooth flow-density relationship, called the *fundamental diagram*.
4. The walking speed will settle at 0 m/s at a specific maximum crowd density.

The first two assumptions were found to be consistent with the data, but the pedestrian-flow theory had to be revised, since the two latter assumptions do not always hold.

When these fundamental parts of pedestrian motion and avoiding maneuvers had been investigated, an improved version of the *social-force-model* [53, 57] was formulated.

In order to enable the revision of previous works and the analysis of the crowd disaster mentioned above, algorithms used for video-tracking have been introduced. The novelty of this work concerns the uniqueness and quantity of data on which the algorithms are validated and calibrated, but also the focus on analyzing millions of pedestrians rather than hundreds.

The aim of this thesis is to move from theoretical models and controlled lab conditions to applicable models for real-world conditions.

Kurzfassung

Als diese Dissertation im Januar 2005 begonnen wurde, nutzten wissenschaftliche Untersuchungen von Fußgängern fast ausschließlich Computersimulationen und Evakuierungsexperimente. Seitdem haben viele Wissenschaftler an einer Verbesserung der Methoden gearbeitet. Heute werden empirische Daten mit Hilfe von Videoanalysen, Laser- und Infrarotkameras erhoben. Jedoch konzentrieren sich viele dieser Arbeiten auf künstliche Setups, in denen sich Fußgängermassen durch Korridore und Engpässe bewegen. Diese Experimente erlauben es, Massenbewegungen zu verstehen. Jedoch gibt es immer noch Forschungslücken.

Es ist schwierig, unter solch kontrollierten Bedingungen Fortschritte darin zu erzielen, die auftretenden Dynamiken bei großen Katastrophen zu verstehen, in denen manchmal Hunderttausende oder sogar Millionen von Fußgängern involviert sind.

Immer wieder kommt es zu Katastrophen in großen Menschenmengen. Leider sind von diesen Ereignissen häufig nur qualitative Informationen anstelle von quantitativen Daten erhältlich. Es ergab sich die besondere Gelegenheit, quantitatives Filmmaterial über eine Katastrophe in Mina (Königreich Saudi-Arabien) zu erhalten. Dort starben am 12. Januar 2006 hunderte von Pilgern während der jährlichen muslimischen Pilgerfahrt nach Mekka.

Mit dem erhobenen Videomaterial konnte nachvollzogen werden, wie die Menschenmenge zuerst unbehindert fließen konnte, dann immer dichter wurde und wie es schließlich zur Katastrophe kam. Von den Erkenntnissen der Analyse der oben beschriebenen Katastrophe konnten neue Methoden entwickelt werden, die dabei helfen können, ähnliche Katastrophen in Zukunft zu vermeiden.

Ein weiterer Beitrag dieser Dissertation besteht darin, einige Annahmen, die üblicherweise bei der Simulation von Fußgängerdynamiken gemacht werden, in Frage zu stellen und zu überarbeiten. Diese Annahmen sind:

1. Ein Fußgänger verhindert Zusammenstöße, indem er seine Schrittgeschwindigkeit so verändert, dass seine Beschleunigung exponentiell mit der Distanz zu dem zu umgehenden Fußgänger oder Objekt abnimmt.

2. Ein Fußgänger zeigt stärkere Reaktionen auf Ereignisse, die vor ihm passieren, als auf Ereignisse, die hinter ihm passieren.
3. Die Bewegung eines in einer Menschenmenge befindlichen Fußgängers folgt immer dem Strömungs–Dichte Verhältnis, was als Fundamental–Diagramm bezeichnet wird.
4. Die Laufgeschwindigkeit eines Fußgängers erreicht bei maximaler Menschendichte einem Wert von 0 m/s .

Die ersten beiden Annahmen wurden von den empirischen Daten bestätigt. Unsere Analysen zeigen jedoch, dass die Annahmen 3 und 4 nicht immer gültig sind. Somit müssen Standardtheorien von Fußgängerdynamiken überarbeitet werden.

Im Anschluß an die Analyse dieser fundamentalen Aspekte von Fußgängerverhalten und dem Verhalten bei Ausweichmanövern wird das Social–Force–Modell [53, 57] weiterentwickelt.

Um auf vorhergehenden Arbeiten aufzubauen und um die oben beschriebene Katastrophe analysieren zu können, werden Algorithmen für die Video–Verfolgung von Fußgängerbewegungen entwickelt. Das Neue bei diesem Teil der Arbeit liegt nicht nur in dem verwendeten Verfahren selbst, sondern auch in der Einzigartigkeit und der großen Menge an verwendeten Daten, die mit diesem Verfahren analysiert werden.

Ein zentrales Ziel dieser Arbeit besteht demnach in einer wissenschaftlichen Weiterentwicklung von theoretischen Modellen und kontrollierten Laborexperimenten hin zu Modellen, die unter realen Bedingungen tatsächlich anwendbar sind.

Die Analyse von Fußgängern ist ein interdisziplinäres Feld, das von verschiedenen wissenschaftlichen Disziplinen mit verschiedenen Zielen betrieben wird. Leider gab es bislang wenig Bemühungen, die Resultate innerhalb dieser Teilgebiete im Rahmen einer konsistenten Theorie zu vereinen. Als seltene Ausnahmen können die Arbeiten von [126] und [10] genannt werden. Diese Dissertation verfolgt das Ziel, diese theoretische Vereinigung weiter voran zu treiben.

Dazu muss man zwischen der Neuerfindung des Rades und der Wiederverwendung nicht geprüfter Resultate abwägen. Dementsprechend ist ein Teil dieser Dissertation dem Vorhaben gewidmet, bisherige Forschung im Lichte empirischer Daten und neuer Methoden zu evaluieren.

Da sich die Arbeit mit recht unterschiedlichen Aspekten von Fußgängerverhalten beschäftigt, konzentriert sich die Analyse in verschiedenen Teilen der Arbeit auf einige

ausgewählte, alternative Modelle. Insbesondere bei der Modellierung und Simulation wird anstelle einer eingehenden Übersicht verschiedener Modelle eine Diskussion des speziellen Social–Force Modells [53, 57, 51] präsentiert. Für Übersichten über verschiedene Simulationsmodelle zu Fußgängerverhalten wird auf die Studien [26] und [52] verwiesen.

Die Dissertation gliedert sich wie folgt. Nach der Einleitung im ersten Kapitel, behandelt Kapitel 2 neue Methoden der Videoanalyse, die sich besonders für Aufzeichnungen von menschenüberfüllten Plätzen eignen. Das dritte Kapitel behandelt die verschiedenen Phänomene der Selbstorganisation in Menschenmengen. Dabei werden neben den bereits bekannten Phänomenen auch zwei neu entdeckte Phänomene, *Stop-and-Go Wellen* und *Crowd Turbulence*, beschrieben. Kapitel 4 erläutert die verschiedenen Aspekte des Fundamentaldiagramms. Wir führen eine neue Herangehensweise ein, bei der anstelle von *lokalen, globale* Quantitäten gemessen werden. Der Fokus liegt dabei, wie in den vorherigen Kapiteln, auf Menschenmengen die eine sehr hohe Dichte aufweisen. Das fünfte Kapitel behandelt das Gebiet der Modellierung und Simulation. Dabei wird eine Methode verwendet, die empirische Daten und Simulationen zusammenführt. Diese Methode bietet den Vorteil, dass zentrale Eigenschaften von Fußgängerinteraktionen extrahiert werden können. Im Anschluß daran werden sowohl die bisherigen als auch die neuen Fußgängermodelle mit Hilfe der empirischen Daten evaluiert. Kapitel 6 stellt dann ein neues Konzept vor, wie die Architektur von Fußgängereinrichtungen automatisch optimiert werden kann, indem evolutionäre Anpassungen in Simulationen vorgenommen werden. Kapitel 7 beschäftigt sich mit der Anwendung der vorher entwickelten Werkzeuge und Modelle, wobei Analysen und Verbesserungen der Sicherheit von Pilgerströmen während der *Hajj* (die jährlich stattfindende muslimische Pilgerung nach Mekka) vorgenommen wurden. In Kapitel 8 wird die vorliegende Arbeit zusammengefasst und zukünftige Forschungsrichtungen aufgezeigt.

Acknowledgements

First, I would like to thank my supervisor Prof. Dr. Dirk Helbing, for providing me with an endless array of challenges on the one hand, and giving me the freedom to explore my own research directions, on the other hand. I am also grateful for his careful reading of the first draft of the thesis, and for his suggestions of improvements.

I would like to express my gratitude to the Ministry of Municipal and Rural Affairs, Kingdom of Saudi Arabia, in particular H. R. H. Prince Miteb bin Abdul Aziz Al Saud and H. R. H. Prince Dr. Mansour bin Miteb bin Abdul Aziz Al Saud for supporting the Hajj-related research, H. E. Dr. Habib Zein Al-Abideen for his generosity in sharing video material from the Hajj and for believing in my work, and Dr. Salim Al-Bosta for his hospitality during my visits in the Kingdom of Saudi Arabia and for his valuable comments on the thesis.

I would also like to thank STESA, especially Andre Mamier, and PTV AG, especially Dr. Peter Vortisch and Stefan Hengst, for giving me the opportunity to test the models and methods within a professional setting, Dr. Tobias Kretz for providing many useful comments and suggestions on the thesis, Prof. Dr. Serge Hoogendoorn at TU Delft for sharing trajectory data from his experiments, Dr. Daniel Nilsson at Lund Institute of Technology for the joint project on pre-movement time in evacuations, Dr. Heiko Rauhut at ETH Zurich and Michael Mäs at the University of Groningen, for translating the summary into German, and the German Research Foundation for funding (DFG project He 2789/7-1).

Further, I would like to thank my colleagues at TU Dresden and ETH Zurich, especially Pradyumn Shukla for his input regarding modeling and optimization, Peter Felten for extracting the pedestrian-trajectory data, Wenjian Yu, Mehdi Moussaid and Karl Walkow for assisting me in the work on modeling and simulation, Dr. Martin Treiber for interesting discussions, Dr. Arne Kesting for sharing his L^AT_EX thesis template, and Sonic Chan for sharing the challenges that come with analyzing huge crowds under special conditions. Finally, I would like to thank my parents Bo and Berit for their support and my wife Emma for her support and love.

1 Introduction

Crowds have been studied for more than a century. For example, Le Bon [90] investigated the crowds during the French Revolution in the 1890s. This rather qualitative study of crowd psychology was later followed by Mintz [97] who conducted an experiment on escape panics, which was later put into a game theoretic framework by Brown [24]. Both Le Bon and Brown pointed out that the collective behavior can not be explained by the average of the actions that the individuals would take individually. Rather, it is something that *emerges*. Still today, these lines of thoughts are relevant and are being investigated with quantitative methods, such as computer simulations and analysis of detailed empirical data from experiments or from real-world observations.

Today, the study of pedestrian crowds is an interdisciplinary field, and is carried out within different communities, with different focuses, e.g.:

- Physicists (or so-called socio-physicists) are focussing on explaining self-organization phenomena in pedestrian crowds.
- Fire engineers are simulating evacuation scenarios, to assess the impact of fires and the building geometries on the evacuation time.
- Biologists and social scientists are investigating how pedestrians interact with each other.
- Computer scientists are setting up large-scale multi-agent simulations (often distributed over many computing nodes), as well as making realistic crowd animations.
- Computer scientists in the field of computer vision use videos of pedestrian movements in order to develop algorithms for identifying and tracking the pedestrians.
- Psychologists are interested in the decision-making mechanisms pedestrians are facing.

- Transportation engineers focus on the mobility of crowds, particularly on the relation of the flow of pedestrians and the crowd density, which is called the *fundamental diagram*.

The list above is of course a crude generalization, since there are many overlaps between the fields.

What can be observed, for example from citation networks and conferences, is that there are many sub-communities within the wide field of pedestrian research. Sadly, there has not been enough effort in connecting the results of all these sub-communities into one consistent theory. There have been a few studies in the past which are working across many of these sub-fields [126, 10], and this thesis is aiming into that direction.

There is a trade-off between re-inventing the wheel and re-using non-validated results. Therefore, a part of this thesis is revising previous research with help of empirical data and tools that did not exist a few years ago.

Since the work is treating a wide range of aspects of pedestrians, part of the thesis is focusing on one of several alternative models. More noticeable, in the part on modeling and simulation, we will focus on the special case of the social-force model [53, 57, 51]. For surveys of alternative simulation models of pedestrian crowds, the reader is encouraged to read Refs. [26, 79, 52].

In favor of the interdisciplinary readability, the terminology in this thesis is kept as general and plain as possible, and does not assume any deep knowledge in any particular field of science.

The thesis is structured as follows: It starts with gathering data, and therefore, Chapter 2 introduces new methods of video analysis, which are able to analyze very crowded videos. Then, Chapter 3 explains different self-organization phenomena in crowds, starting with an introduction of previously known phenomena and then introduces two newly found phenomena, i.e. *stop-and-go waves* and *crowd turbulence*. Chapter 4 covers various aspects of the fundamental diagram, with an introduction of a new way to measure *local* rather than *global* quantities. As in the previous chapter, there is a focus on high crowd densities. Chapter 5 treats the subject of modeling and simulation, by using a method of fusing empirical data with simulations. This method gives the possibility to extract various properties of pedestrian interactions. Then, previous pedestrian models as well as improvements of them are evaluated. Chapter 6 introduces a new concept of automatically optimizing the architecture of pedestrian facilities, by evolutionary adjustment of simulations. Chapter 7 contains applications of the tools and models previously developed, to analyze and improve the safety of pilgrim flows, during the Hajj, i.e. the

annual Muslim pilgrimage to Makkah. Chapter 8 concludes the work and presents future research directions.

2 Video Analysis

2.1 Introduction

As in all fields of science, empirical evidence must eventually be found, that supports the theories. So also in pedestrian research. In the past, a popular way of empirically validating pedestrian models has been to design the model in a way that it can reproduce as many of the observed (self-organization) phenomena as possible (see Chapter 3). There are also some popular ways of quantitatively validating the simulation models, that have been frequently used in the past, e.g.:

1. comparing the fundamental diagram obtained from empirical data, to the fundamental diagram obtained from experiments [95],
2. comparing cumulative outflows from evacuation simulations with similar curves, obtained from evacuation experiments [88, 84].

A great source of information about the interactions and behaviors of pedestrians can be found in video recordings. The recordings can be used in different ways, with an increasing amount of detail and effort:

1. Visual inspection of the videos, to get a better (qualitative) understanding of the general mechanisms of how pedestrians interact, e.g. to see if the pedestrians tend to overtake each other, organize themselves into lanes, physically interact with each other, and so on.
2. Extraction of macroscopic quantities by manually counting the number of people moving into different parts of the systems [51]. This method can determine crowd densities, flow rates and velocities.
3. Automatic or semi-automatic video analysis performed by video-tracking software [126, 69, 80].

Note that, besides video analysis, there are also other methods to extract pedestrian trajectories, e.g. using infra-red sensors [83] or using laser devices [144].

There are various ways to classify videos. They can be analog or digital, recorded or a (live) stream, and they can be encoded in many different ways. However, the only thing the reader has to know by now is that a video is composed of a series of images, called *frames*, which are separated with a fixed time interval Δt .

Depending on which level of detail is needed for the extracted data, two different approaches are presented here. One extracts the complete paths (trajectories) of all pedestrians and the other method extracts aggregated values. The first approach is named the *Microscopic Approach* and is a semi-automatic approach, which means that computer software is attempting to track the pedestrians, but where it fails, the user can manually make corrections during the tracking process (by using the computer mouse). Also, the computer will *not* attempt to identify all the heads. Rather the user has to click on all new heads once they appear in the video, and these heads will then be tracked until they leave the boundaries of the video. With this approach, one can reach 100% accuracy (in terms of no missed or interchanged trajectories), if the user is careful and the video is not extremely crowded (to the extent that the user can not clearly identify all the heads). This approach is sensible if the video contains up to dozens of pedestrians appearing simultaneously, and the total length of the video is on the time scale of minutes.

The second approach, called the *Macroscopic Approach* is used for larger areas with more pedestrians and a higher level of crowd density (pedestrians per m²), where complete trajectories are not needed but rather macroscopic, aggregated quantities like crowd density, flow rate and velocity.

The two different video-analysis approaches have different objectives,

- *Microscopic Approach:* This approach is typically used to extract trajectories in videos with not more than hundreds of pedestrians in total, and over a few minutes only. As a consequence, the tracking can be much slower than real time¹. Tracking a few minutes of a video recording, with a total of 100 pedestrians can take an hour. But what is lost in time is gained in accuracy.
- *Macroscopic Approach:* This approach is used to analyze larger areas, where up to thousands of pedestrians appear simultaneously and on time scales of days or

¹With *real time*, we mean that we are able to process several frames per second, and therefore not have any gaps in our collected data (when processing live video streams).

even longer. This approach is designed to run in real time and needs no human interaction during operation.

The motivation for using two different approaches is that for microscopic validation and calibration of microscopic models, reliable trajectory data are needed. On the other hand, for online² monitoring of crowds, it is a requirement that the algorithm can process the videos in real-time. If the algorithm can not process the videos in real-time, it means that one must post-process the data, which further implies that the results can not be used for direct crowd control. Another useful application of the *Macroscopic Approach* is that the video analysis can be made on arbitrarily long timescales of days or even longer, since no human interaction is needed during operation.

Note that, using video-analysis to identify and track pedestrians has been performed by various other people before. However, there were various reasons to develop our own video-analysis algorithms. It was difficult to find any commercial or free software that was flexible enough for our needs. Most available video-tracking software had one or several of the following limitations:

- The software was tailored towards security applications and gave alarm if pedestrians (or pedestrian-movement patterns) were being identified in certain zones, but there was no functionality to analyze the motion of the pedestrians, once they were identified.
- The software was designed to simultaneously track up to dozens of pedestrians, but not hundreds of pedestrians.
- There was no possibility of interacting with the software and plugging in self-developed modules, to investigate new kinds of measurements.
- The software was not automated enough. We had to analyze Terabytes (see Chapter 7) of video material, distributed on hundreds of video files. Obviously, it would not be possible to manually load each file and set up and execute a video-analysis run on each of these files. Rather there was a need for an automated software to run different batches of video-analysis on several computers.

Now, after having given a motivation for developing new video-analysis software, we will give an overview of each of the two video-analysis approaches and explain the algorithms behind them.

²The word “online” should not be interpreted in the sense of “the Internet”, but rather to distinguish *live processing* from *post-processing*.

However, before introducing our methods, the state-of-the art in identifying and tracking pedestrians is presented.

2.1.1 State-of-the Art

The field of computer vision is very active in research related to the identification, classification and tracking of objects from videos, but the aims are very different. Most studies use fixed cameras, but some also use mobile cameras, for example in the automobile industry, to automatically identify pedestrians at the sides of the road and give alarms upon detection [111, 141].

The fields of image analysis and pattern recognition have developed accurate methods to identify objects from images. However, with the additional temporal information obtained from series of images in videos, one can gain more accuracy in the identification process of objects.

Video-tracking is (at least) a two-step process; first *identification* of objects and then *tracking* of these objects. Good reviews of video-tracking of pedestrians can be found in Refs. [74, 10, 43].

It has been reported in Ref. [134] that the number of people (and crowd density) has a linear relationship both with the number of edge pixels (15% error for 15 pedestrians) as well as the number of foreground³ pixels (6% error for 15 pedestrians).

However, the relationship is probably not linear for higher crowd densities, because of occlusion, but it can still be used as a reference and give a fast crowd-density estimate as input to other more sophisticated methods.

Identification of pedestrians is often made by first subtracting the background of the video, to separate the background from the foreground, and then apply morphological filters which give a number of “blobs”⁴ as result. These blobs can be investigated by numerous methods to classify them either as pedestrians or as other objects. If the crowd density is low and there is not much interaction between pedestrians, one can simply track these blobs directly [73].

However, more sophisticated methods are often applied in order to classify the blobs,

³The foreground pixels are defined as the pixels which are more different than a certain threshold from the background, and the background can be determined from a frame without any people in it. The edge pixels on the other hand, are defined as the pixels identified with an edge-finding algorithm and suitable thresholding.

⁴These “blobs” can represent anything from parts of pedestrians/objects to several close pedestrians/objects belonging to the same blob.

i.e. Artificial Neural Networks (ANN), Support Vector Machines (SVM) [141], or shape-based methods [37].

Some authors consider human bodies as rigid objects and therefore match blobs with models representing different body parts [129, 140]. Even 3D-skeleton models of pedestrians are sometimes used.

Contour tracking is another method to recognize humans. The contour can be approximated by “snakes” [82] or by Bsplines [12, 19].

Instead of using a single identification method, it is often more effective to extract several different features and combine them for the identification. For example, the study in Ref. [107] uses Haar wavelets, applied in three different orientations and on different scales. The responses of the wavelet filters, are then used as *feature vectors*. Other popular features include pose, color and texture.

When the feature vectors are calculated, the next step is to use them to classify different objects. Often Bayesian reasoning is used to find the most likely class for a given feature vector. The aim is to find a model with parameters X in frame Z , which tends to be a very complicated task. Instead, one can find \hat{X} , which is the maximum-likelihood estimate of X . Since the posterior distribution $p(X|Z)$ is difficult to calculate, Bayes’ theorem is applied,

$$p(X|Z) = \frac{p(Z|X)p(X)}{p(Z)} \quad (2.1)$$

where $p(Z)$ is a normalizing constant and $p(Z|X)$ is considered to be a function of X for a fixed Z and can be estimated from real data. $p(X)$ is called the prior probability.

Temporal information can also be used to identify humans. Human walking is characterized by a periodic motion with a frequency of about 2 steps per second. Therefore, the Fast Fourier Transform (FFT) or the Discrete Cosine Transform (DCT) could be applied to the trajectory of an object to determine if the object is a walking human or something different.

Other popular approaches to pedestrian tracking include “*optical flow*” which is segmenting areas of consistent flow into objects [89, 18], or identifying changes in pose over time [23].

Nowadays, many pedestrian trackers are using probabilistic methods to estimate the state of each pedestrian, depending on the history (trajectory). For this purpose, the Kalman Filter or the Extended Kalman Filter for non-linear processes, is often used [45, 75]. Recently, Particle Filters [71] have taken over the role of (extended) Kalman filters. Both these methods are special cases of the framework of *Bayesian Filtering*.

When pedestrians are interacting with each other, the tracking process becomes more difficult because of changes in velocity as well as the possible confusion of two close pedestrians. If the density is too high, the pedestrian movements are complicated to predict and the Kalman filters are not as accurate anymore. Another similar method worth mentioning is the *Probabilistic Data Association Filter* (PDAF) and the improved JPDAF (Joint PDAF) [13, 123] which is a sub-optimal approximation to the optimal Bayesian filter.

Principal Components Analysis (PCA) from multi-variate statistics can be used for tasks such as face recognition or recognition of lip motion, but it is sometimes also applied to identify pedestrians [74]. PCA is used to reduce multidimensional data sets to lower dimensions, and once the number of dimensions has been reduced, other methods can be applied to the resulting lower-dimensional data.

Often, not only pedestrians need to be identified or tracked, but also vehicles [93]. However, this is not within the scope of this thesis.

2.2 Microscopic Approach

With the *Microscopic Approach*, the aim is to determine accurate trajectories of the pedestrians from crowd videos. To reach this goal, the following pragmatic approach is used: The computer attempts to automatically track the pedestrians, but where it fails, the user has to correct for the error. The tracking procedure works as follows:

1. A video file is loaded.
2. The user selects an origin and an x and a y axis.
3. The user marks the diameter of a head, for reference.
4. The user clicks at each new (unmarked) head on the screen.
5. The user presses a button to progress the video one timestep, Δt , where Δt is typically about 1/8 s.
6. The computer tracks all heads from the previous frame to the new frame.
7. For the heads that the computer did not manage to track (this happens typically for about 5% of the heads), the user corrects for the errors by moving the displaced points with the computer mouse.

8. If the tracking is not finished, go to step 5.

Assume that for the frame, F_t at time t , $n(t)$ heads have been identified, at locations $\vec{r}_i(t)$ with $i = 1..n(t)$. Now the task of the algorithm is to track each head from one frame to the next one. This is done by determining the head displacement $\Delta\vec{r}_i(t) = \vec{r}_i(t + \Delta t) - \vec{r}_i(t)$ by comparing the pixel values within the head radius R from one frame to the next one, minimizing the deviation

$$\sum_{||\vec{x}|| \leq R} |F_t[\vec{r}_i(t) + \vec{x}] - F_{t+\Delta t}[\vec{r}_i(t) + \Delta\vec{r}_i(t) + \vec{x}]| \quad (2.2)$$

by variation of $\Delta\vec{r}_i(t)$ (within a realistic range of possible locations). The matching is made within a disc of size R , corresponding to the size of a head.

The search is limited by a maximum searching distance D reflecting the maximum walking speed $v_{max} = 4$ m/s, i.e. $D = v_{max}\Delta t$.

The reason why heads are tracked and not the feet, or the center-of-mass projected on the floor, or something else, is simply because the head is the clearest visible part of the body, from the videos.

After the tracking is made, there are also a few post-processing steps that must be applied.

The first post-processing step is to compensate for the distortion, resulting when a spherical scene is being projected onto a flat surface. To compensate for the distortion, we use an algorithm based on bilinear interpolation, described in [41].

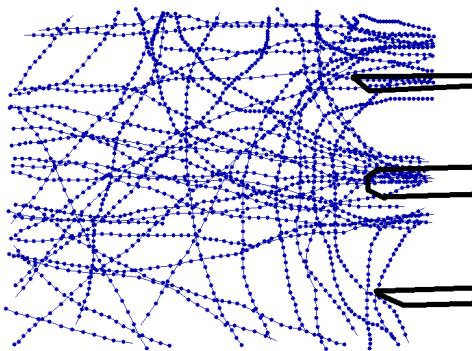
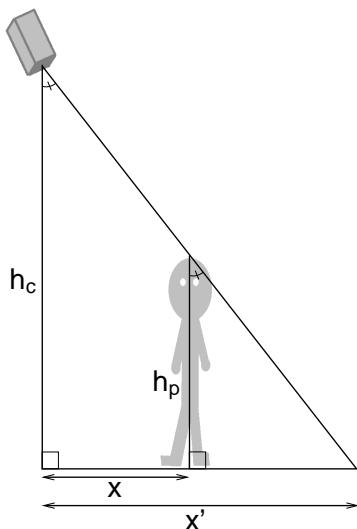
Second, we are interested in the pedestrians' center-of-masses projected on the floor, rather than the head positions. To obtain these values, we need to know the distance from the ground to the camera and the height of the people. Since we do not have explicit information about the height of the people in the videos, we assume that all persons have a height of 1.7 m.

Given the height h_c of the camera, the height h_p of the person, and the measured head position x' , the transformation onto the floor is obtained by using the fact that the two triangles in Fig. 2.1 are similar. Therefore,

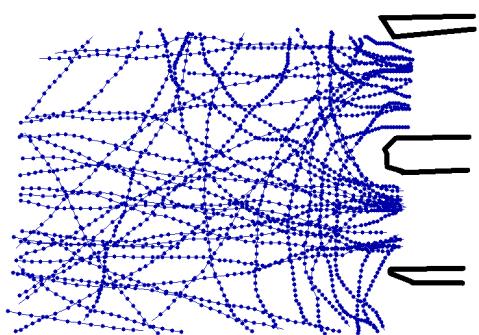
$$\frac{x'}{h_c} = \frac{x' - x}{h_p} \Rightarrow x = x' \frac{h_c - h_p}{h_c} \quad (2.3)$$

The same transform is also applied on the y coordinates.

Note that, according to Eq. (2.3), the compensation is smaller, the larger the camera height h_c is. Therefore, it is advantageous to mount the camera as high as possible.



Original data



Transformed data

Figure 2.1: Video tracking, used to extract the trajectories of pedestrians from video recordings at the end of two escalators in Budapest, Hungary. Top left: A snapshot of the application being used to track the movement of a few pedestrians. Note that the floor tiling can be used as a reference coordinate system. Top right: Illustration of how to transform the pedestrian head coordinates into floor coordinates. Bottom left: The raw data of the resulting trajectories, i.e. the head positions from the video. Bottom right: The trajectories after being transformed onto the ground plane. Note that the pedestrian positions fit much better to the position of the escalators after the transformations have been applied.

Mounting the camera higher will, as a side effect, also decrease the distortion, because it is not required to use a wide-angle lens. However, it is even more important to put the camera pointing perpendicular to the ground. Large deviations from the perpendicular mounting will result in more complicated transforms to obtain the floor coordinates. If the video camera is mounted high enough and perpendicular to the ground, the only reference needed is the size of the area on the ground being monitored.

2.3 Macroscopic Approach

2.3.1 Introduction

The semi-automatic video-analysis method presented in Sec. 2.2 will now be extended to a fully automated video-analysis method. Note, however, that this full automatization only holds when the video-analysis is in operation. It still needs some manual user interference in the initial calibration phase.

One extension of the video-analysis method presented in this section is that the method will not only track the pedestrians, but it will also identify new persons once they enter the video.

The idea behind the software is that it will automatically determine heads by searching for round structures in the video. For this purpose, several digital filters (transformations) are sequentially applied to the video frames to identify the relevant structures. Finally, we obtain the positions $\vec{r}_i(t)$ and speeds $\vec{v}_i(t) = d\vec{r}_i/dt$ of all pedestrians i in the recorded area as a function of time.

For each frame, the heads are located at positions $\vec{r}_i(t)$. By comparing these positions with their local neighborhoods in the next video frame, one can determine the velocity $\vec{v}_i(t)$ of the pedestrians.

Table 2.1 shows the simplified core of the video-analysis method. To gain more accuracy, the following steps are added:

- Double-checking of the identified circles by applying an Artificial Neural Network trained to recognize heads.
- Application of statistical compensations obtained by extensive manual counts.

In the following sections, the details of the video-analysis method are given.

2.3.2 Hough Transform

The input is a grayscale frame F_{ij} , which initially consists of three color channels, red (R), green (G) and blue (B), and is created as $F_{ij} = (R_{ij} + G_{ij} + B_{ij})/3$. The index i denotes the vertical position, while j denotes the horizontal position of each of the pixels in the frame.

As calibration data we have the samples $\{F, \vec{r}_i\}_j$, containing a list of frames, F , and a list of head positions \vec{r}_i in each of these frames j . The head positions are manually determined, by clicking with the computer mouse at all of the heads i within each of the training frames j . We are using three different datasets (see Fig. 2.2).

Step	Operation	Illustration	Formula
	Original frame		$F^{0'}$
#1	Lens correction		$F^0 = \mathcal{L}(F^{0'})$
#2	Edge detection (Sobel)		$F^S = \mathcal{S}(F^0)$
#3	Thresholding		$F^T = \mathcal{T}(F^S, T')$
#4	Hough transform		$F^H = \mathcal{T}(F^T, r')$
#5	Extract heads (circles)		$\vec{r}_i = \mathcal{C}(F^H, \gamma'), i \in [1, n_h]$

Table 2.1: The chain of transforms carried out, in order to obtain the locations of the human heads from a video frame $F^{0'}$. There are three parameters r' , T' and γ' , which need to be calibrated for each individual video or camera. See Tab. 2.2 for the meaning of these parameters.

For recognizing patterns, Artificial Neural Networks (ANN) have proven to be a good tool. However, since we need to process the video stream in real time, scanning for heads at every single location in the frame takes far too long time. Therefore, we make use of the Hough transform [70], to find circular structures in the frame, corresponding to the heads of the pedestrians. When we have a list of head-position estimates, we employ the ANN to double-check these positions. By using this scheme, we combine the computational speed of the Hough transform with the accuracy of the ANN. This scheme greatly reduces the required frequency of applying the ANN, and the task can now be performed in real time.

The first thing that has to be done is to identify the objects in the frame. This is done by finding the edges in the frame, which can be done by the Sobel operator $F^S = \mathcal{S}(F)$ [121]. Then, the edges are transformed into binary values $F^T = \mathcal{T}(F^S, T')$ by using a threshold function,

$$\mathcal{T}(F_{ij}, T') = \begin{cases} 0 & \text{if } F_{ij} \leq T' \\ 1 & \text{otherwise.} \end{cases}, \text{ for all } i, j. \quad (2.4)$$

The Hough Transform [70] is then applied to get the probability surface corresponding

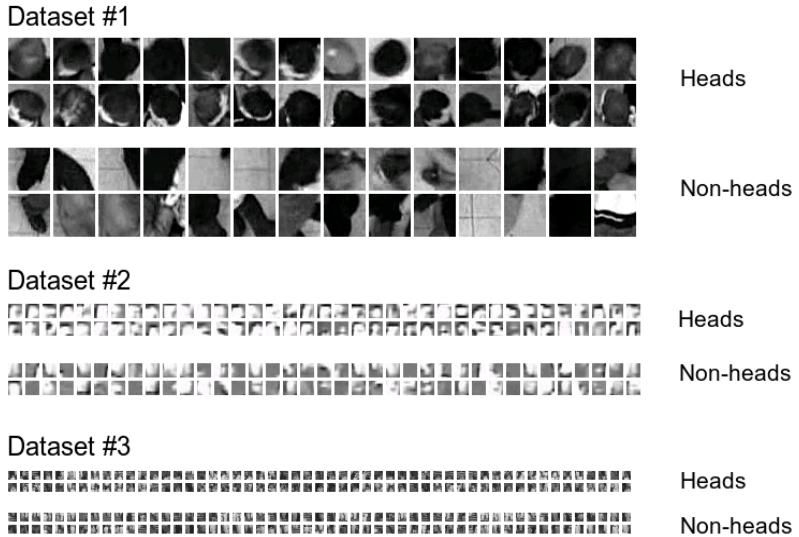


Figure 2.2: Samples from the three different datasets. Dataset #1 consists of 11 frames and 196 patterns (heads and non-heads), dataset #2 of 5 frames and 255 patterns, and dataset #3 consists of 3 frames and 240 patterns. The heads and the non-heads are representing 50% of the samples each.

Parameter	Meaning
r'	The mean head radius.
T'	The edge threshold, separating the pixels from the Sobel frame as <i>edge pixels</i> and <i>non-edge pixels</i> .
γ'	The fraction of heads (Hough threshold), defining how large fraction of the head circumference that must be identified in order for the observation to pass as a head.

Table 2.2: A summary of the three free parameters and their meanings. The typical head radius r' can easily be determined as the average head radius from a series of head samples. The two remaining values are a bit more complicated to determine, but can be found by an optimization algorithm, which is minimizing the error between automatic and manually determined densities or flows.

to the likeliness of finding a head at each position in the frame, $F^H = \mathcal{H}(F^T, r')$.

Finally, the list of all head locations is obtained by, $\vec{r}_k = \mathcal{C}(F^H, \gamma')$, $k \in [1, n_h]$, defined as,

$$\mathcal{C}(F^H, \gamma') = \text{all points } \{i, j\} \text{ that satisfy: } \begin{cases} F_{ij}^H \text{ is local maximum within radius } r' \\ F_{ij}^H \geq \gamma' \end{cases} \quad (2.5)$$

The parameter γ' defines what fraction of the head circumference must be identified in order for the observation to pass as a head.

The complete chain of computation to obtain the head positions \vec{r}_i from the frame F^0 can be found in Tab. 2.1.

In the following sections, we will refer to the chain of transforms in Tab. 2.1 as *the Hough method*. Figure 2.3 shows an example for each of the transforms.

Details on the three parameters r' , T' and γ' are given in Tab. 2.2.

2.3.3 Artificial Neural Network

Next, we set up an Artificial Neural Network (ANN), to recognize the heads. For positive samples (heads), we use our positions \vec{r}_i from the training data, but for negative samples (non-heads), we have to be more careful. By picking random locations not closer than 0.1 m from any of the heads \vec{r}_i from the training data, the ANN will easily learn to recognize heads, since most of the “non-heads” correspond to patches of asphalt or

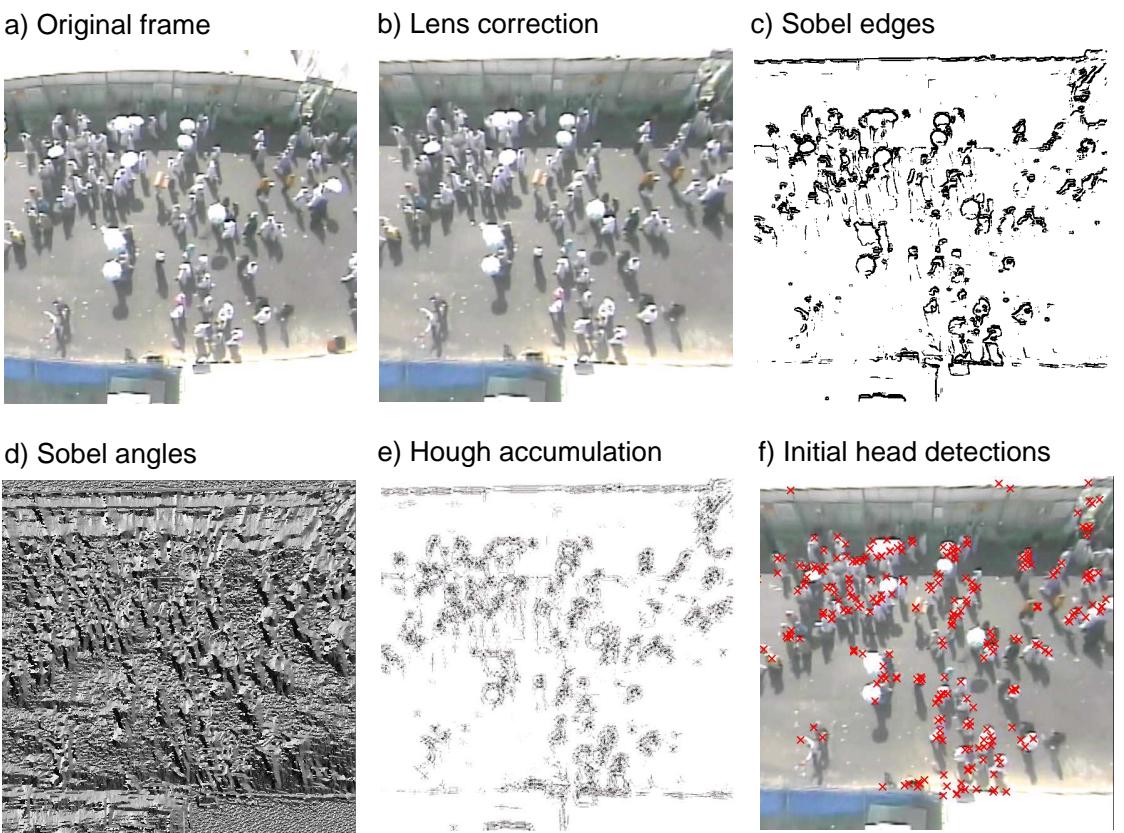


Figure 2.3: Sample frames for different transforms. (a) Original frame F^0' . (b) The frame F^0 , after lens correction has been applied. (c) The Sobel edges, after the edge detection has been applied. (d) The angles, from the Sobel edge-detection algorithm. (e) Frame F^H , after the Hough algorithm has been applied. (f) The initial estimation of head locations \vec{r}_i , when thresholding has been applied on F^H according to Eq. (2.5). Note that these head locations correspond to the first estimate, before the ANN has been applied, which then separates heads from non-heads (see Fig. 2.9).

other trivial patterns with very little similarity to human heads. Therefore, when the ANN is given the task to double-check the heads identified by the Hough method (see Tab. 2.1), the accuracy is not much increased, since most of the non-head samples are patches of asphalt and other trivial things, easy to distinguish from heads. For this reason, we generate negative samples (non-heads) from those points that the Hough method mistakenly picked out as a head, but are in fact *not* heads according to the training data. Now, the ANN gets a more difficult (and realistic) task - to differentiate

between heads and structures looking similar to heads.

The Artificial Neural Network (ANN) that we use, is a feed-forward, backpropagating network with multiple hidden layers. The layout of the network is influenced by the

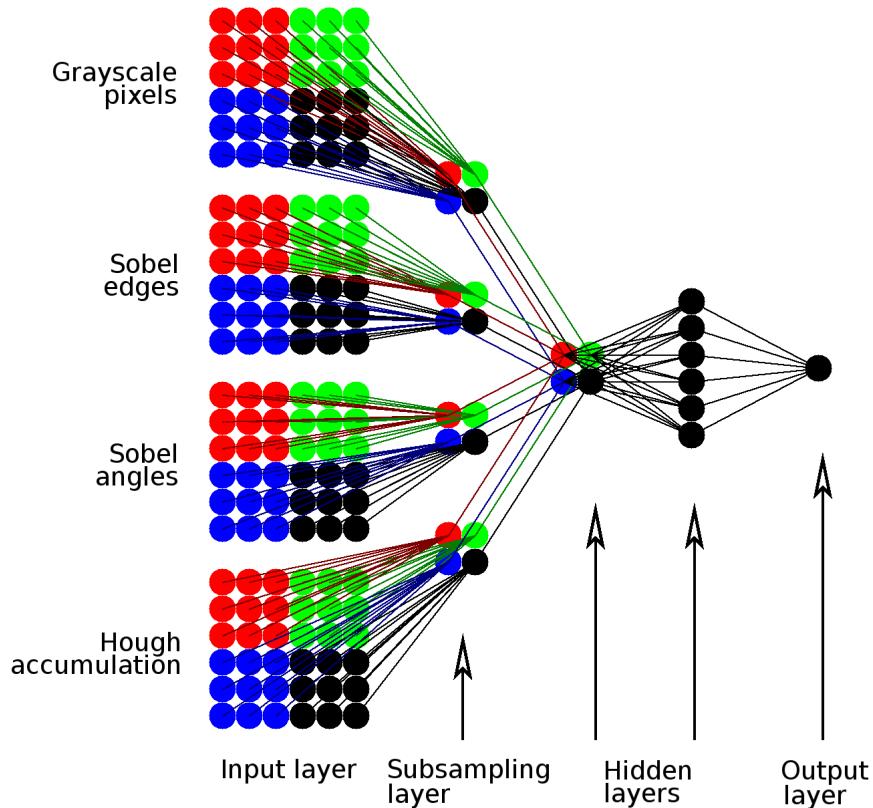


Figure 2.4: The topology of our Artificial Neural Network (ANN), used to classify heads, where the processing is done from left to right. The input is a combination of grayscale pixel data, edge magnitudes from the Sobel transform, tangential angles from the Sobel transform, and the Hough accumulation data. The input is then sub-sampled to reduce the complexity. Note that the network is *not* fully connected. Rather, the inputs from similar geometric locations are clustered together. After the sub-sampling, all different sub-sampled inputs are merged into one layer, and finally a last layer of fully connected nodes are used for the classification. Note that this figure is only to illustrate how the topology looks like. In reality, the input consists of more than only 6×6 pixels, which gives more nodes also in the two following layers.

Convolutional Neural Network [91, 106], used to classify handwritten characters. However, we do not perform the convolutional operations like edge-finding in the ANN itself, but rather as a pre-processing step, to save computational time. The idea behind the structure of our ANN is to create an ANN with the following properties,

- The accuracy must be as high as possible.
- Several different inputs can be combined, like grayscale pixels, edge magnitudes, edge-tangent angles and Hough-accumulation values.
- Different kinds of information originating from the same location in the input frame are processed together in the same nodes.
- Input data are with highest possible resolution, but are then sub-sampled, to reduce the complexity (computational effort) of the ANN.
- Finally, the last hidden layer in the network contains a variable number of nodes which can make the network adaptive to the complexity of the problem.

The topology of the ANN (see Fig. 2.4) consists of an *input layer*, where nearby pixels are sub-sampled and processed together in the same node in the *sub-sampling layer*. Then, data from different inputs (grayscale pixels, Sobel edges, Sobel angles and Hough accumulation) are merged, where data from the same locations are processed together in the same node in the first *hidden layer*. The last hidden layer consists of N hidden nodes, which are fully connected to the single output node. Finally, there is one last node in the *output layer* which will give 0 for *no head* or 1 for *head*.

We are using different inputs, because applying all of them together will give higher accuracy than using only one of them. The reason for the sub-sampling is to make the ANN less sensitive to small variations in rotation, scaling and translation of the input data. This does not mean that the ANN is fully invariant to any of the three transformations (rotation, scaling or translation), but it is robust to small deviations in any of them.

The input to the ANN is picked from a box within the video frame, with a size of $L \times L$ m² (which implies an increasing number of nodes the higher the resolution of the video).

Fine-tuning of the ANN

Artificial Neural Networks (ANNs) can be a very powerful tool, if properly designed. Wrongly designed, however, they may give very bad results. Since there is no frame-

work/rule on how to determine the optimal design of an ANN for a given problem, we will now investigate how the classification error depends on the design of the ANN. We have five frames with altogether 129 heads and 129 non-heads (dataset #2 in Fig. 2.2) that we split into two parts - one training set and one validation set with 50% heads and 50% non-heads each. Note that there is no overlap between the two datasets. The idea is to train the ANN on the training dataset and validate the ANN on the validation dataset. We increase the number of hidden nodes as long as the error for the validation data is decreasing, but at a certain point, the ANN will be overtrained, *remembering* patterns rather than *recognizing* them. This point is clearly visible, since the two errors will diverge, i.e. the training error will decrease while the validation error will *increase*.

When introducing even more hidden nodes, a second transition will occur, where the ANN is too complex to converge at all, resulting in maximum errors for both the training-set as well as the validation set. Figure 2.5 shows the errors, as a function of the number of hidden nodes, for three different sizes of the input-data kernels, $L = 0.28$ m, 0.38 m, and 0.48 m.

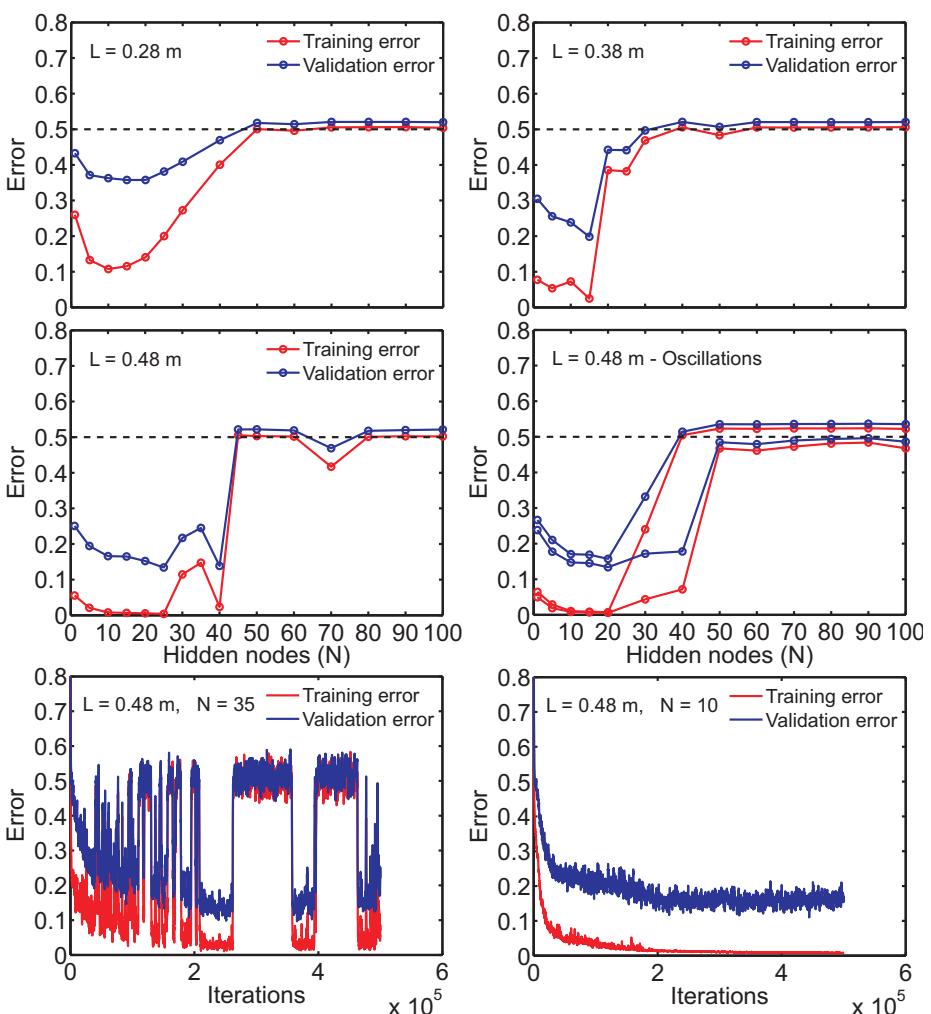


Figure 2.5: Errors of the classification by the Artificial Neural Network (ANN), for the kernel size $L = 0.28\text{ m}$ (top left), $L = 0.38\text{ m}$ (top right) and for $L = 0.48\text{ m}$ (middle left). In this example, $N = 20$ and $L = 0.48\text{ m}$ gives a training error of almost zero and a validation error of 15%. Middle right: Results when splitting each error curve into two parts – the average error of the 50% highest error values as well as average error for the 50% lowest values. When the curves are separated from each other it means that the error curve is not converging to a certain value, as it normally does. Instead, the error is oscillating between two different error levels. Bottom left: When the number of hidden nodes is $N = 35$, the error does not converge. Instead it is oscillating between two different levels. Bottom right: The time-evolution of the error level is converging for $N = 10$ hidden nodes. Note that there is no universal value of the number of hidden nodes that would work perfectly for all videos. Rather the number of hidden nodes is a free parameter to be calibrated for each video.

2.3.4 Principal Components Analysis

As the results in Sec. 2.3.3 have shown, a higher number of hidden nodes in the ANN leads to a decreasing error for a while, but after a certain number of hidden nodes, the error is *increasing* again. The reason for this is partly due to the limited size of the training dataset. It would be tempting to spend time on collecting more data, in order to push the error lower. However, this is only possible to a certain extent. The patterns used in Sec. 2.3.3, have a size of around 20×20 pixels. Each pixel can take a number of values, giving altogether a parameter space of 400 dimensions. With this in mind, it is clear that no matter how much effort is spent on manually collecting empirical data, the dataset will always sparsely populate the space spanned by these 400 dimensions. This is a well known problem in the field of classification and is referred to as *The curse of dimensionality* [28].

Luckily, there are ways to decrease the number of dimensions. One way is to use *Principal Components Analysis (PCA)* [17], which is a linear orthogonal transformation used to project a high-dimensional dataset onto a lower dimensional coordinate system. The result is a reorganization of the data, where the first coordinate is the one “explaining” the largest variance of the data, the second coordinate explains the second largest amount of the variance of the data, and so on. After performing PCA, we can drop most dimensions, since the relevant information is predominantly captured by the lower dimensions.

The PCA involves the following steps:

Preparation, given a collection of N patterns of size $n \times m$:

1. Linearize each pattern by putting it, row by row, into a vector with $d = n \times m$ elements. Put all these linearized patterns into a matrix A , with d columns and N rows. Now, the matrix A represents one training pattern by each row and one dimension by each column.
2. Compute the mean $\langle A_j \rangle$ for each column j . Subtract $\langle A_j \rangle$ from A to obtain a matrix \hat{A} , where each column has the sum 0.

Principal Components Analysis (PCA):

3. Determine the covariance matrix $C = cov(\hat{A})$.
4. Calculate the eigenvalues λ_i and the eigenvectors \vec{x}_i for C .

5. Sort the pairs $\{\lambda_i, \vec{x}_i\}$ by decreasing values of λ_i .
6. Create a new coordinate system via the k orthogonal eigenvectors \vec{x}_i with the k largest eigenvalues λ_i . This spans a k -dimensional space.
7. The coordinates of an N -dimensional pattern P , along these k coordinates, represent the pattern in the reduced k -dimensional system.

The PCA transform detailed above, is a linear transform which is sometimes also referred to as the *Karhunen-Loëve transform*. Alternatively, one can use non-linear PCA specifications, for example the *Independent Component Analysis (ICA)* [17].

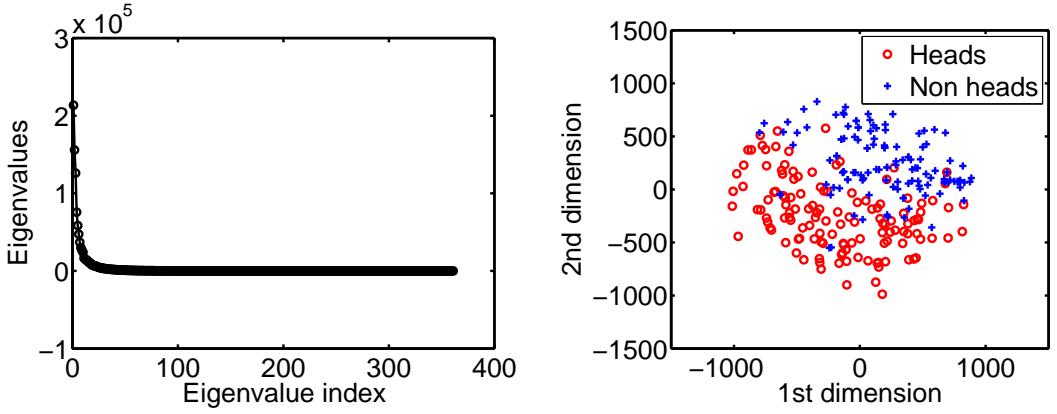


Figure 2.6: Left: Eigenvalues from the PCA, ordered by value. Right: Sample data projected onto the two-dimensional space spanned by the first and second principal axes determined by the PCA.

Let us now apply the PCA, as detailed above, to a dataset containing 258 patterns (dataset #2 in Fig. 2.2), with 50% heads and 50% non-heads, respectively. By performing the PCA and plotting the sorted eigenvalues we clearly see (Fig. 2.6) that most of the information is contained only within a few dimensions. If we carry out a projection of the samples onto only the first two dimensions, \vec{x}_1 and \vec{x}_2 , we can clearly observe the clustering of the patterns into areas representing heads and non-heads (see Fig. 2.6).

Since the samples corresponding to the two classes (heads vs. non-heads) are (at least partly) separable, the task is now to create a *decision surface*, which makes an optimal cut between the two classes. The easiest choice would be to use a hyperplane to separate the two classes, and this hyperplane could be found by a least-squares method. Another common method to optimize the cut with a hyperplane is to use Support Vector

Machines (SVM) [141, 17]. In order not to complicate it further and to keep the amount of computation low, we chose again to use our Artificial Neural Network (ANN) to find a decision surface. Note that one clear advantage of using an ANN is that the decision surface is not restricted to a linear hyperplane, but can be a complex surface (where the degree of complexity is increasing with the number and layers of hidden nodes).

See Fig. 2.7 for classification of the patterns after the dimension has been reduced with PCA. Two different classification methods are used, a linear least-squares method and an ANN with 10 hidden nodes. Generally, the more of the primary components are used, the better the linear least-squares method works. The ANN on the other hand works best when only a few of the primary components are used and the accuracy is decreasing for higher numbers of primary components.

2.3.5 Making Use of the Temporal Information in Videos

Now, after we have developed methods for identifying pedestrian locations and for determining their velocities (Eq. (2.2)), the last step is to put these methods together and analyze videos rather than single frames. Until now, we did not make use of the richer amount of information that exists in a series of frames, i.e. the temporal information in videos. If a pedestrian is identified at time t , at position $\vec{x}(t)$, and with velocity $\vec{v}(t)$, we know from the mechanics of human walking, that this very same pedestrian is likely to be found somewhere close to the position $\vec{x}(t) + \vec{v}(t)\Delta t$ at time $t + \Delta t$, with a precision that is decreasing with an increasing forecasting time Δt . To test this hypothesis, we use three trajectory datasets obtained by the method from Sec. 2.2 applied on the three videos in Fig. 5.2. For each of these three datasets, for each of the pedestrians, and for each starting time t_0 , we determine the distance error $\|\vec{x}(t_0) + \vec{v}(t_0)\Delta t - \vec{x}(t_0 + \Delta t)\|$, between the linearly extrapolated position $\vec{x}(t_0) + \vec{v}(t_0)\Delta t$ and the actual position $\vec{x}(t_0 + \Delta t)$. We now analyze the distributions of distance errors for the different forecasting times $\Delta t = 0.5, 1.5, 2.5, 3.5$ and 4.5 seconds. Figure 2.8(left) shows how the probability-density function of the distance errors evolve by a growing forcecasting time Δt . By averaging the distance error, as a function of the forecasting time Δt , we get a relation that is approximately linear for videos B and C, but super linear for video A (see Fig. 2.8(right)). The lower level of predictability for video A can be understood by the fact that the pedestrians can choose between five destinations (using two different escalators, walking into the left direction, walking above the escalators, or walking below the escalators). Video B, on the other hand, is recorded above a free floor, which makes it very likely that the pedestrians will follow their current directions, without making any sudden

changes in direction or speed. For the experimental setup in video C, the pedestrians are instructed to follow certain paths (walking into the left direction or walking into the right direction), which makes also this scene more predictable than video A.

As mentioned in Sec. 2.1.1, there are various methods based on Bayesian Filtering, for example the Kalman Filter [45, 75] and the Particle Filter [71]. These are recursive methods that assign a “model” to each tracked object, and from the history of observations of the same object, they draw conclusions on how these objects will most likely evolve over time.

The disadvantage of this kind of methods is that they need considerable computational efforts, especially if a large number of objects are to be tracked at the same time. For post-processing of videos, the computational effort may not pose a serious problem. However, if the videos are processed in real time, a crowded scene results in longer processing time, which results in a longer time elapsing between the frames, which worsens the accuracy of tracking, which leads to errors in tracking, and eventually the situation may get out of hand.

For this reason, we choose not to keep track of the evolution of single individuals over longer distances, but we still want to make use of the temporal information in the videos, to increase the accuracy of the tracking. From Fig. 2.8(left), we see that the diffusion of a pedestrian from his/her linearly extrapolated path, is approximately growing linearly with the forecasting time Δt . Therefore, if we assume that a head detection at position $\vec{x}(t)$ at time t gives probability p that there is a real head at this position, we can represent this with a Gaussian distribution with its center in $\vec{x}(t)$ and with a standard deviation $\sigma \rightarrow 0$ (i.e. Dirac’s delta). This Gaussian distribution will then evolve in time with its center at position $\vec{x}(t) + \vec{v}(t)\Delta t$ and its standard deviation $\sigma = k\Delta t$, where k is a fit parameter.

An example frame where the probability-density-surface of head detections is overlaid onto the original frame is shown in Fig. 2.9. The most likely head detections are marked by crosses.

2.3.6 Parameter Calibration

Given the calibration data $\{F, \vec{r}_i\}_j$ (see Fig. 2.2), the average head radius can be extracted by applying the Hough Transform [70], for all r' in the range 0 m to 0.25 m. Figure 2.10 depicts the average value of the Hough transform, as a function of r' . We see that this function has a peak at $r' = 0.1$ m, which is the value of the head radius we will use in the following.

The task to find the parameters T' , γ' remains. As a first step, to see how these two parameters are influencing the measurements, error curves are plotted as a function of T' for 10 different values of γ' (see Fig. 2.11).

The error is defined as,

$$E = \left\langle \frac{Q_a - Q_m}{Q_m} \right\rangle \quad (2.6)$$

where Q_a is the flow rate (during 1 second) which is automatically determined by the algorithm, and Q_m is the flow rate from the calibration data, which is obtained by manually clicking the heads during 1 second.

We define $E(T', \gamma')$ as the error given by Eq. (2.6), where the value of Q_a is obtained by video-analysis performed with the parameters T' and γ' .

According to Fig. 2.11, there are several combinations of T', γ' that minimize the error ($E = 0$). Notice the steeper slopes of the error-curves, when only the Hough method is used, compared to when we apply the ANN classifier as a second step. A low slope is a desired property, since the error would then be less sensitive to noise.

We now define an optimization algorithm, which automatically finds suitable values of the edge-threshold parameter T' as well as the fraction-of-head parameter γ' . See Tab. 2.2 for a detailed description of these parameters. The objective function is to finding the set of parameters, minimizing the absolute value of the error $|E|$, as defined by Eq. (2.6).

However, since there are many equally good solutions ($E = 0$), we restrict the optimization function to find a solution in the interval $0.25 \leq T' \leq 0.75$. The algorithm is based on two nested blocks of binary search (see Algorithm 1). If we intend to find the single best parameter set, we would have to include one more objective in our optimization algorithm (except minimizing $|E|$). For example, we could, from all points where $E = 0$, choose the one which has the lowest slope dE/dT' , of the error. However, for now we stick to the simpler objective function.

2.3.7 Evaluation of Accuracy

In order to assess the accuracy of the video-tracking method, the flow is measured during 72 hours and compared to the flow estimated from manual counting (see Chapter 7). It turns out that the agreement of the two curves is much better when applying the second classification step with the ANN, compared to only using the Hough method for the classification (see Fig. 2.12). The videos used for the comparison were recorded in a

Algorithm 1 Optimization Algorithm

```
 $\gamma_{min} \leftarrow 0$ 
 $\gamma_{max} \leftarrow 1$ 
 $success \leftarrow \text{false}$ 
while  $success = \text{false}$  do
     $\gamma' \leftarrow (\gamma_{min} + \gamma_{max})/2$ 
     $T_{min} \leftarrow 0$ 
     $T_{max} \leftarrow 1$ 
    while  $T_{max} - T_{min} > 0.01$  do
         $T' \leftarrow (T_{min} + T_{max})/2$ 
        if  $E(T', \gamma') > 0$  then
             $T_{min} \leftarrow T'$ 
        else
             $T_{max} \leftarrow T'$ 
        end if
    end while
    if  $T' < 0.25$  then
         $\gamma_{min} \leftarrow \gamma'$ 
    else if  $T' > 0.75$  then
         $\gamma_{max} \leftarrow \gamma'$ 
    else
         $success \leftarrow \text{true}$ 
    end if
end while
```

very crowded street (see Fig 3.2 in Sec. 3.6) in Mina, Kingdom of Saudi Arabia, during the Pilgrimage to Makkah [54]. As a reference, the following scheme was carried out:

- The 72-hour-video material was played on a large TV screen at 1/10th of the original speed.
- During 5 seconds, two persons were counting the number of persons passing a central line in the video, separately for both directions.
- This procedure was repeated every 10 minutes for the full 72-hour period, giving 432 flow-rate samples for each direction and for each of the two persons.

The mean of the samples from the two persons were used as a reference⁵.

The samples were then interpolated to get an estimation of the flow rate as well as the total number of pedestrians passing the street during the 72-hour period. When observing the total number of pedestrians counted within 72 hours, this number matches fairly good when using the Hough method only. However, when observing the flow rates, we see that there are large fluctuations. When employing the ANN as a second step though, we see that also the flow rates are matching much better. Note that there will always be some remaining uncertainty in this method because of the sampling method. Therefore we can not say exactly, how accurate the algorithms are, but what we can clearly observe is the relative accuracy of our two methods. The videos correspond to the ones used for producing dataset #2 in Fig. 2.2.

See Sec. 7.5 for a thorough validation of the software, on a real-life application.

⁵The inaccuracy of the manual counts is negligible as compared to the inaccuracy due to the sampling of the time series (counting over 5 seconds, every 10 minutes). The reason is that the crowd density is extremely high only during very short peak times. Most of the time, the crowd density is on a level that allows one to distinguish all pilgrims in the videos.

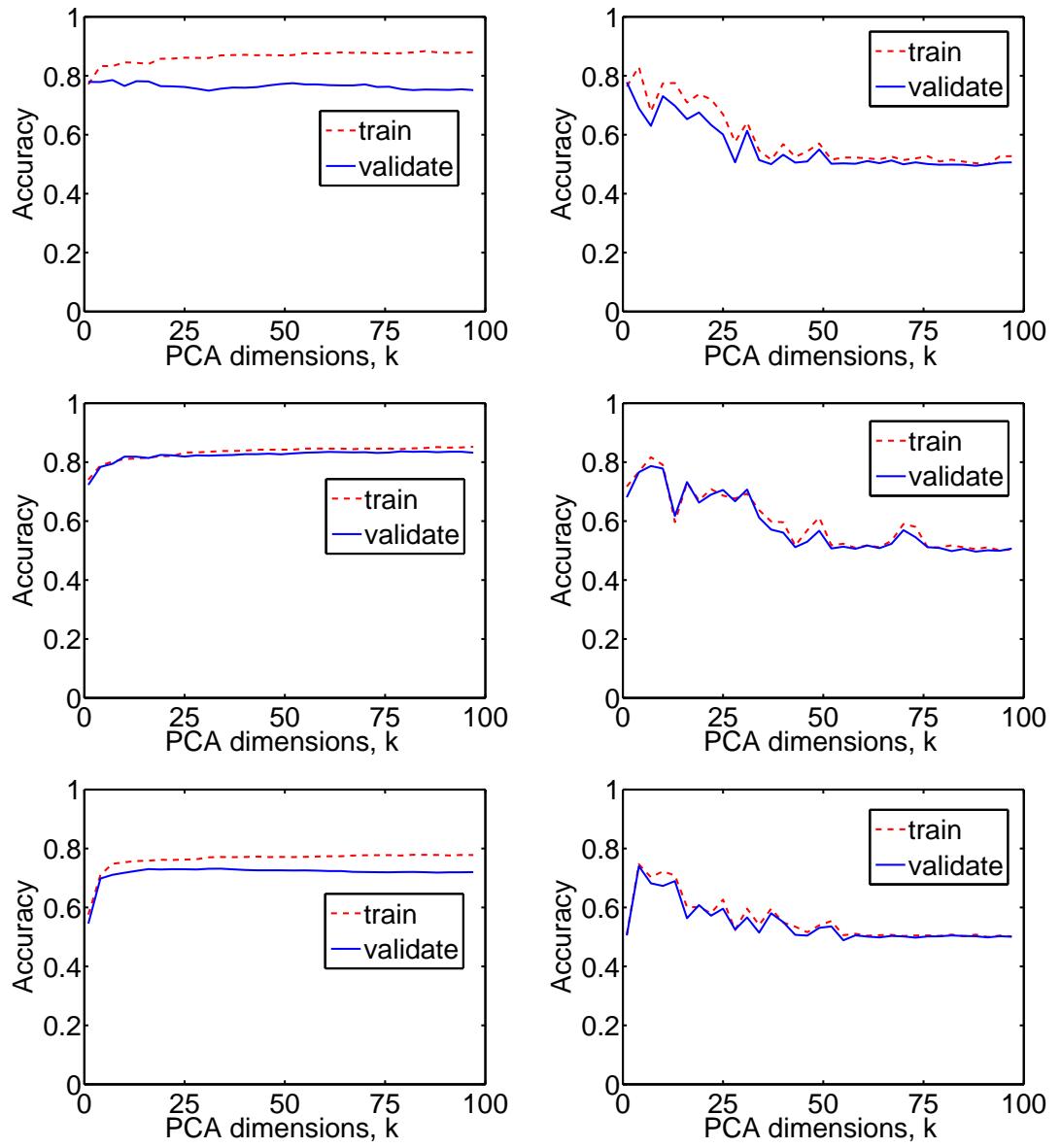


Figure 2.7: Classifications of the data projected onto the k primary components, found by the PCA method. The results are shown for three different datasets, dataset #1 (top), dataset #2 (middle) and dataset #3 (bottom). Left: Classification by fitting a hyperplane with the least-squares method. Right: Classification by finding the decision surface with an ANN. The accuracy is defined as the fraction of patterns which are correctly classified as heads vs. non-heads. There is no universal k value that can always be used. Rather, k is a parameter that can be modified, either to increase computational speed, or to increase the accuracy.

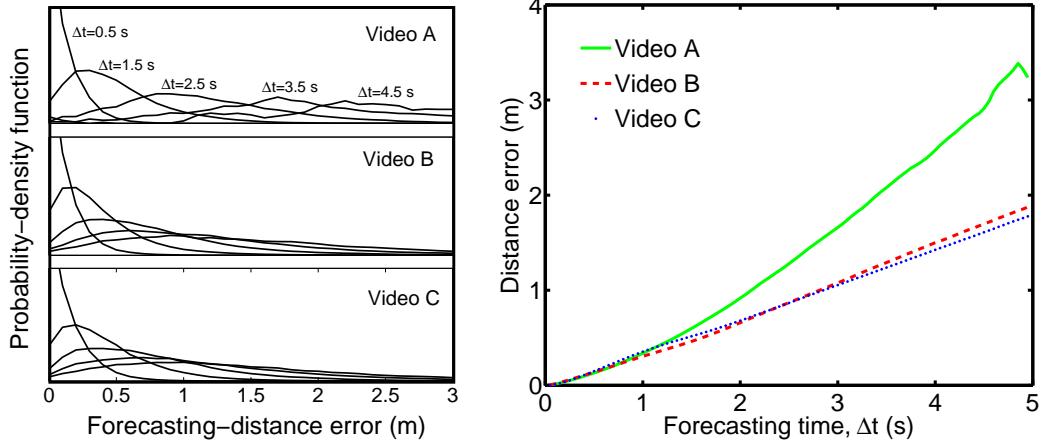


Figure 2.8: For three trajectory datasets (called videos A, B, and C), the distance error between extrapolated positions and real positions after a forecasting time of Δt seconds is determined. Left: Probability-density functions for the forecasting error are determined for different values of Δt . Right: The mean distance error as a function of the forecasting time Δt . The error by linearly extrapolate the pedestrian positions is approximately linear for the videos B and C, and super linear for video A.

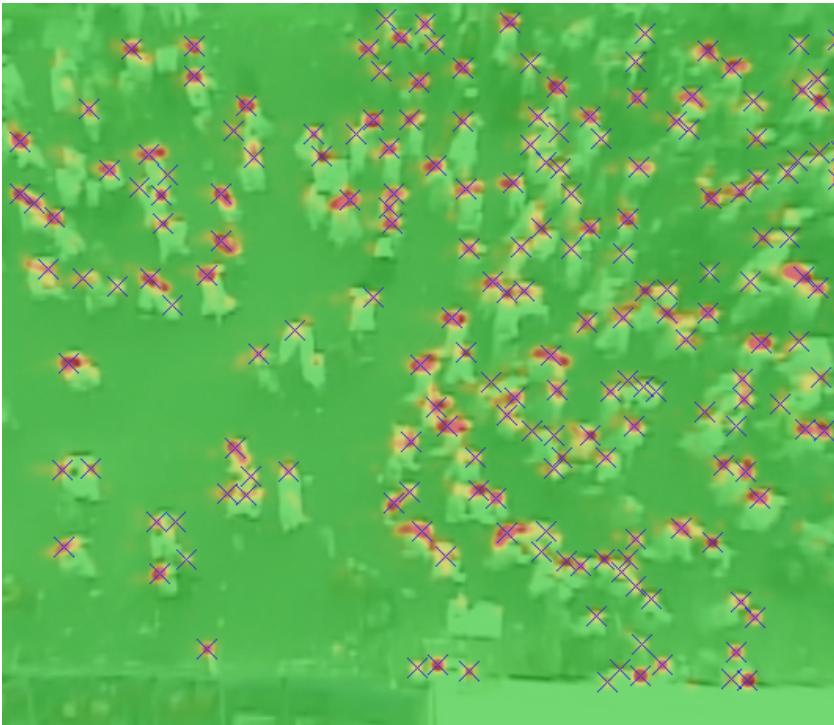


Figure 2.9: Probability-density surface of head detections, overlaid onto the original frame. The crosses show the most likely head locations, obtained by Eq. (2.5), but with the ANN applied as a second step, to improve the accuracy of head detections. Also, not only one frame, but a series of frames have been used to improve the accuracy further. As compared to Fig. 2.3 f, the number of false negatives have been significantly decreased.

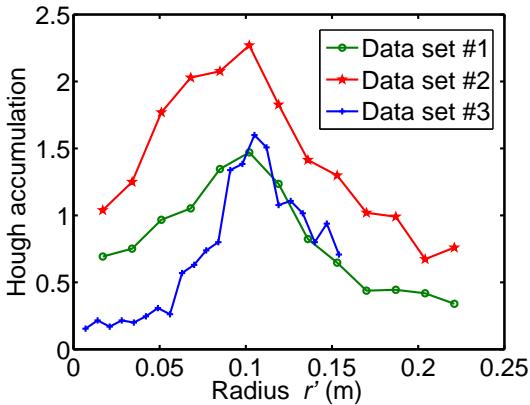


Figure 2.10: The average head radius is found by locating the peak of the average value of the Hough transform, as a function of the radius parameter r' . The datasets for the three curves are shown in Fig. 2.2. Note that the head radius agrees for all three datasets.

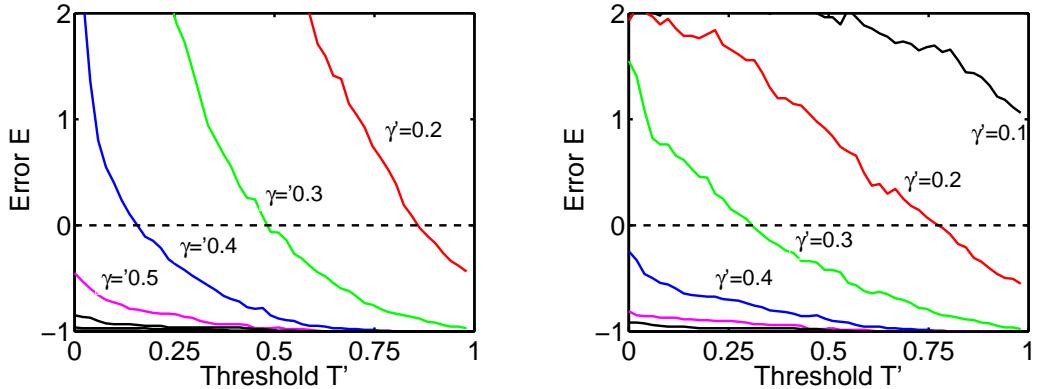


Figure 2.11: The error as a function of the threshold, T' , for different values of the fraction-of-head parameter γ' . Left: Only the Hough method is used for head classification. Right: The Hough method + the ANN method are used for head classification. One can see that the slope is reduced for the curves where the ANN is used as a second step in the classification, compared to when only the Hough algorithm is used. This implies that the second approach is less sensitive to noise.

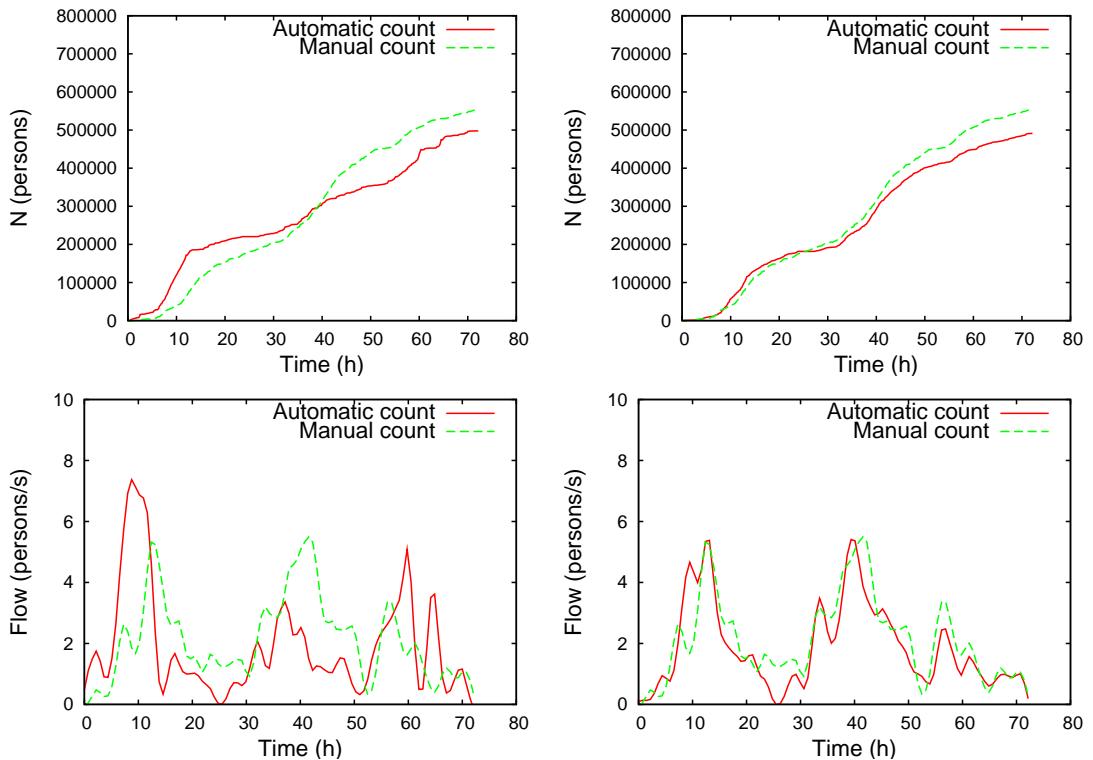


Figure 2.12: Comparison of the cumulative counts and flow rates of pedestrians according to our algorithm as compared to an estimation made by manually counting pedestrians for a short time period every 10 minutes. Top left: Cumulative counts made with the Hough method. Top right: Cumulative counts made with the Hough method and subsequent application of an artificial neural network (ANN). Bottom left: Flow rates computed with the Hough method. Bottom right: Flow rates computed with the Hough method combined with the ANN method. Note that, when adding the ANN method as a second step, there is not a large improvement in the total count of people after 72 hours. The flow rate, however, is much better estimated when applying the ANN method as a second step, as compared to using only the Hough method.

3 Self-Organization Phenomena

3.1 Introduction

Given a set of simple rules of local interactions between pedestrians, complex dynamical patterns can emerge in the crowd. These collective spatio-temporal patterns of motion are referred to as self-organization phenomena.

In the following sections, a survey will be provided on work that has been done in the past. In Secs. 3.6 and 3.7, two newly discovered self-organization phenomena will be presented, which are only observed in crowds with very high crowd densities.

3.2 Lane Formation

If the crowd density is low, slight changes in walking speed are enough to avoid potential encounters with the few pedestrians around. However, when the crowd density is higher and especially when there is a variation of desired directions of walking, the avoidance manoeuvres become more complex. Let us start with the the special case when there are two streams of oppositely moving pedestrians, intersecting at an 180-degrees angle, i.e. bi-directional movement. From empirical observations of such situations [51], it has been reported that the pedestrians self-organize into lanes of uniform direction of motion. This organization of lanes is beneficial for the pedestrians, since it reduces the frequency and strength of avoidance manoeuvres. There is nothing that indicates that the organization into lanes follows from an individual or collective planning. For a majority of the pedestrians, it is doubtful if they are even aware of the collective pattern that they are part of. As it turns out, a very simple model, where pedestrians are “pushed” aside (either physically or psychologically) from oppositely moving pedestrians can reproduce the lane formation phenomenon.

3.3 Stripe Formation

If we generalize the principle of lane formation to two pedestrian flows intersecting at an arbitrary angle, different from 180 degrees, it turns out that self-organization occurs also in this case, increasing the efficiency of walking for all pedestrians. This case calls for more complex avoidance manoeuvres than for bi-directional streams, since the pedestrians do not only have to avoid collisions with oppositely moving pedestrians, but they eventually also have to traverse the stream of oppositely moving pedestrians. It turns out that the resulting dynamics of the pedestrians reminds of the zipper strategy observed in vehicular traffic when two lanes are merging into one. Pedestrians with the same desired walking direction turn out to form groups, and the walkway is dynamically subdivided into stripes made up by groups with the two different directions of motion. This phenomenon is called *stripe formation*.

3.4 Pedestrian Trail Formation

Another pedestrian self-organization phenomenon is trail formation [48, 56]. Analogously to the phenomenon of lane formation described in Sec. 3.2, for movement in terrain like grass or snow, it is easier to follow the steps of others, rather than producing new paths. This behavior is similar to the pheromone-based trail following of ants [31, 32, 109], but for pedestrians the role of attractive chemicals is replaced by the greater comfort of a trail that is used more often. However, if the path of a predecessor deviates too much from the desired path of the pedestrian, he or she will generate a new path in the terrain. When these two principles are put together with some added stochasticity, permanent paths will eventually evolve in the terrain. Interestingly, the shape of these trails looks differently from conventionally engineered pedestrian walkways. Especially when it comes to the angle of two paths merging or separating, naturally evolving paths will usually not have the 90-degree angle of a “T” shaped intersection. Pedestrians rather prefer “Y”-shaped intersection designs, which is common for freeways.

3.5 Bottlenecks

3.5.1 Uni-Directional Bottleneck Flows

Naturally, the most problematic locations during evacuations are bottlenecks, i.e. the most narrow places. The traditional view of a bottleneck is that the flow of pedestrians

is proportional to the walkable width and that the flow characteristics are not further changed. This is roughly true for large widths, but for small bottlenecks, like doors, things become more complex.

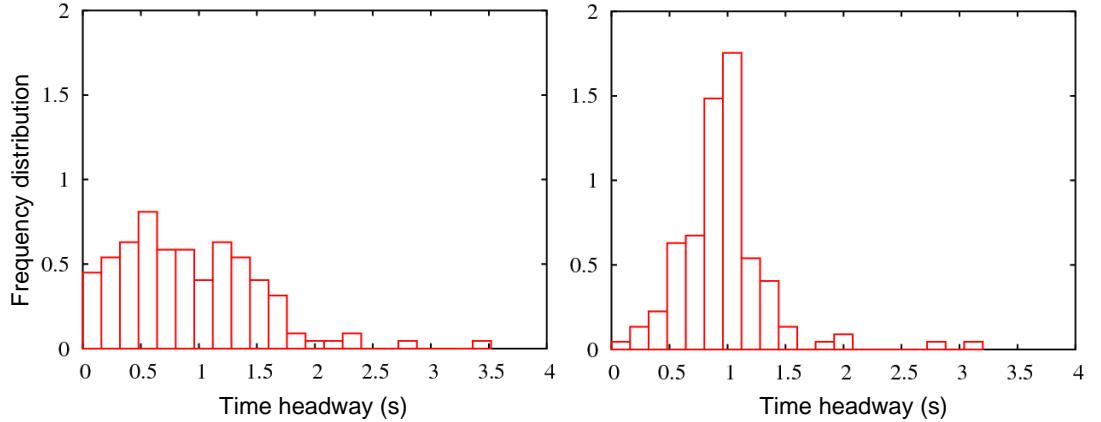


Figure 3.1: Time headway distributions of pedestrians in a uni-directional stream before a short bottleneck (left) and after it (right). The mean values of the inverse time gaps (the flows) are the same before and after the bottleneck. However, due to the interactions in the bottleneck area, the probability-density function is narrower and the maximum is more pronounced after the bottleneck (after Ref. [51]).

For example, at narrow bottlenecks, it is not possible anymore for pedestrians to overtake their predecessors. This removes one degree of freedom. Experiments have shown that the distribution of time headways is more distinct directly after a narrow bottleneck than before, as Fig. 3.1 shows. This indicates that narrow bottlenecks are homogenizing pedestrian flows. See Refs. [66, 87, 118, 51] for further experimental studies of bottlenecks.

Intermittent flows

In situations when there is a jam in front of a narrow bottleneck, e.g. jamming in front of exits during evacuations, incoordination leads to intermittent flows [55]. These intermittent flows are characterized by an alternation between bursts of smooth outflow and times when nobody passes through the door. Intermittent flows are a big concern during evacuations, since they give very unpredictable outflows. During evacuations,

people want to leave the building as fast as possible, making them more impatient and uncoordinated, which gives raise to the “faster is slower” effect [53], i.e. an overall decrease in the outflow as a result of a stronger urge to leave the building. This is sometimes referred to as *escape panics*, and was experimentally demonstrated by Mintz [97]. In his experiment, a number of cones were put in a bottle where water was pouring into the bottle from the bottom. A string was attached to each of the cones and the ends of the strings were given to the participants. A participant got rewarded if his/her cone came out dry, and got increasingly fined when there was water in the cone. The outcome was that there was always a jamming of cones at the bottleneck, if the water rose too fast. Intermittent flows and how to deal with them will be treated in Sec. 6.3.

3.5.2 Bi-Directional Bottleneck Flows

So far, only uni-directional flows through bottlenecks have been discussed, but for the case of bi-directional flows, another interesting self-organization phenomenon occurs. When people are using a bottleneck in both directions and the bottleneck is narrow enough to allow only one person to pass at a time, it is obvious that the flow can either be zero or in one of the two directions. But for wider bottlenecks, we could have bidirectional flows. However, for crowded scenarios this happens only for wide bottlenecks, and not immediately when the bottleneck is wide enough to let two persons pass each other.

Once the flow is established in one of the two directions, this will generate a forward momentum that sustains the flow for some time. However, the crowd that wants to pass in the other direction will increase in number and they will eventually be eager enough to stop the flow, and change the flow direction, so they can pass the bottleneck. As this scheme is repeated, the resulting flow is oscillatory in time.

3.6 Stop-and-Go Waves

Until now, we have seen that, when the crowd density is increasing, pedestrians are coordinating themselves with the people around, which results in self-organization.

When the density is so high that coordination is difficult, will self-organization survive or will it break down as well?

We have discovered that, when the density is so high that the level of coordination breaks down, the smooth flow, turns into stop-and-go flow [54]. For stop-and-go flow, rather than everybody moving at the same speed, the motion is characterized by an

alternation of moving and stopping. Basically, the pedestrians are stopped until free space appears in front of them and they can then take one or several steps forward. Looking at the phenomenon from a macroscopic level, waves of free space are propagating upstreams in the crowd, anti-parallel to the movement of the pedestrians.

By analyzing video-recordings where stop-and-go waves have been observed in reality (see Fig. 3.2) it turns out that stop-and-go waves move with an almost constant propagation speed. Details of the measurement site and the analysis method are given in Chapter 7.

3.7 Crowd Turbulence

When the crowd exhibits stop-and-go flows and the crowd density is increased even further, there is a second transition, at which the motion becomes turbulent and is characterized by blocks of people being moved in all possible directions (see Fig. 3.3).

In order to compare the three flow regimes, laminar (smooth) flow, stop-and-go flow and turbulent flow, Fig. 3.4 depicts three typical trajectories, belonging to each of the three regimes. The position x is the distance to the entrance of the Jamarat Bridge (see Chapter 7) to which the pedestrians are moving.

The time is scaled to make the comparison more clear, and there is a symbol every 5 seconds to show the time scale. Note that the stop-and-go trajectory is characterized by a step function, with partly vertical lines corresponding to the stopped time period and sudden diagonal shifts corresponding to a wave. The trajectory from the turbulence regime is characterized by motion both into the positive as well as into the negative x -direction.

During high crowd densities, the mechanical pressure onto the people in the crowd is high, but it is not homogeneous. This conclusion can be drawn from the considerable spatial variation of densities, even for very large average densities (see Fig. 4.1). From this inhomogeneity of forces, eruptive events of pressure release may follow. During these events, people experience large (involuntary) displacements, which cause an irregular crowd motion into all possible directions. By measuring the distance between two subsequent stops (defined as a walking speed below 0.1 m/s) for thousands of pedestrians, the resulting probability-density function over the distances, reminds of a power law distribution. Its exponent is -2.01 ± 0.15 (see Fig. 3.4). This distribution is, to some extent, analogous to the Gutenberg-Richter law [39], which describes the frequency of earthquake amplitudes.

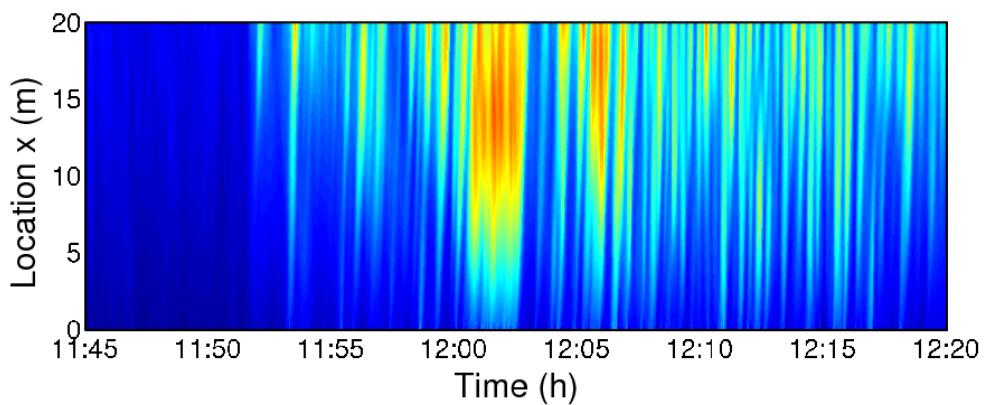
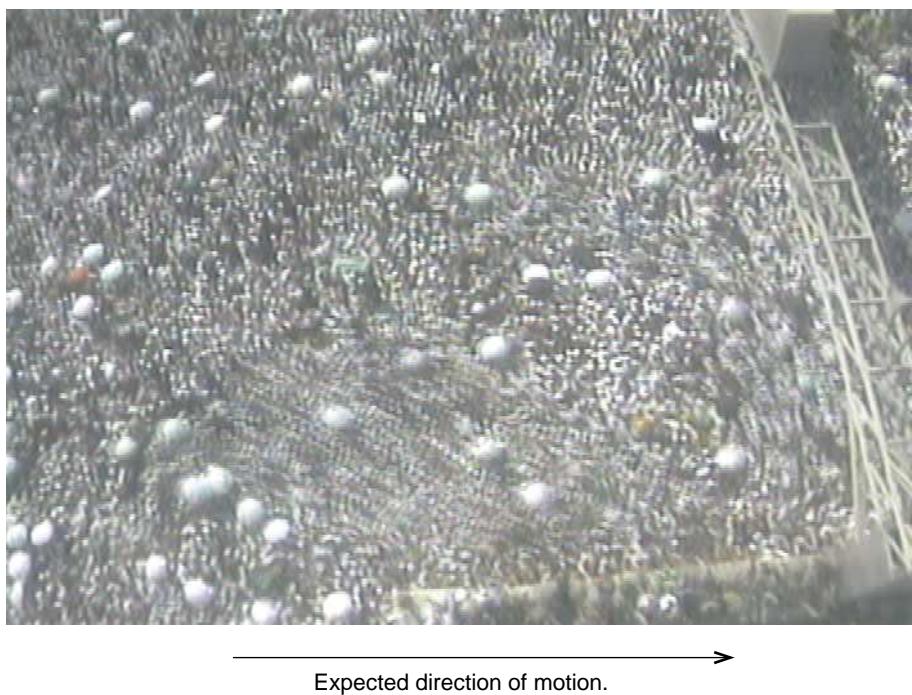


Figure 3.2: Illustration of stop-and-go waves. Top: A long-term photograph of the pilgrim crowds in Mina, Kingdom of Saudi Arabia, in 2006. Blurred parts correspond to areas where people are moving, sharp parts to areas where people are stopped. The wave front corresponds to the line separating these two areas. Bottom: By measuring the density as a function of time and location, one can visualize the transition from free flow to stop-and-go waves around 11:53 AM, on January 12, 2006, at the entrance to the Jamarat Bridge. Once the transition has occurred, the stop-and-go waves persist for a long time period. Blue corresponds to low density and red corresponds to high density, and yellow to intermediate densities.



Expected direction of motion.

Figure 3.3: Long-term photograph of a pedestrian crowd in Mina, Kingdom of Saudi Arabia, during the pilgrimage in January 2006. Crowd turbulence is characterized by blocks of people moving into various directions (blurred areas).

As we do not observe large eddies, the similarity with fluid turbulence¹ is limited, but there is an analogy to the turbulence at currency exchange markets [38].

The reason for calling this phenomenon “*crowd turbulence*” [54] is that the motion visually reminds of turbulence. Moreover, by measuring certain quantities one can show that crowd turbulence exhibits properties that are similar to those of turbulent flows.

For example, the *structure function* [29] $D(\Delta r) = \langle [v(r, t) - v(r + \Delta r, t)]^2 \rangle$ as a function of the displacement Δr has turbulent properties (see Fig 3.5).

Moreover, when measuring the velocity change $\vec{v}(t + \tau) - \vec{v}(t)$ over a time period τ at a given location, and averaging over different locations \vec{r} and times t , we find the results shown in Fig. 3.5. The probability-density functions are calculated for different values of τ both in the laminar-flow regime as well as in the turbulent regime. For small values of τ in the turbulent regime, the velocity probability-density function is sharply peaked rather than Gaussian, which is typical for turbulence.

¹Note that comparisons of pedestrian crowds and fluid dynamics have been made before [72, 61, 47].

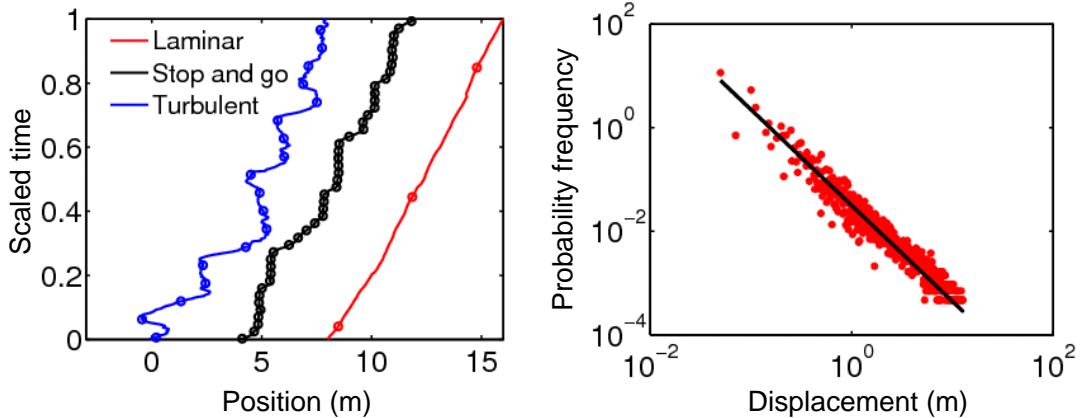


Figure 3.4: Left: Typical trajectories for laminar (smooth) flow, stop-and-go flow, and turbulent flow. Each symbol is separated by 5 seconds. Right: Measuring the distance between two subsequent stops of thousands of pedestrians and making a histogram over the distances reveals a distribution which reminds of a power law with exponent -2.01 ± 0.15 .

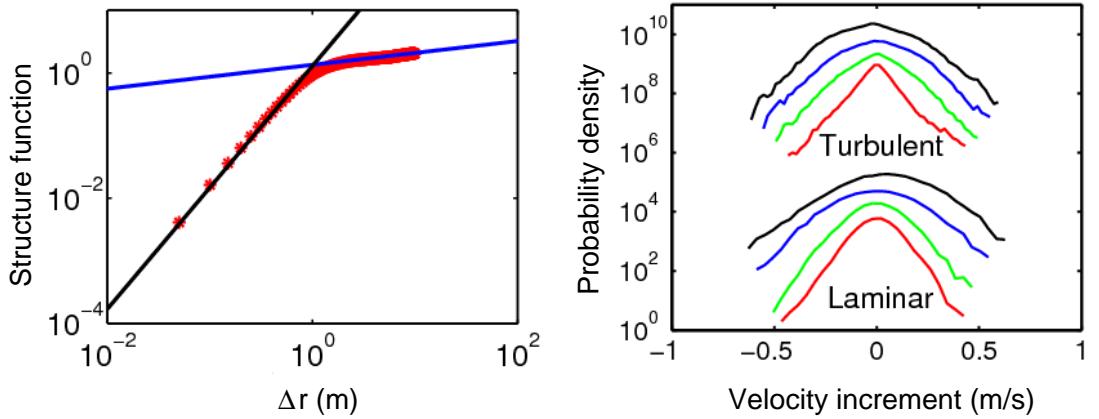


Figure 3.5: Analysis of laminar as compared to turbulent flows. Left: Structure function $D(\Delta r) = \langle [v(r, t) - v(r + \Delta r, t)]^2 \rangle$ as a function of the displacement Δr . As in fluids, the slope at small distances is 2.0, and at large distances, there is another slope (0.18 ± 0.02) because of the (increased) self-propulsion [143] of pedestrians. In contrast to molecules, people are trying to gain more space, in dense crowds. Right: Histograms of velocity increments $\vec{v}(t + \tau) - \vec{v}(t)$. The curves from top to bottom corresponds to a time shift τ of 100s, 10s, 1s and 0.1s. Note that the curves are displaced in the vertical direction to make the figure more clear.

4 Flow-Density Diagram

4.1 Introduction

The fundamental diagram [62, 116, 117, 136, 81] is a relation for how the specific flow (in unit $m^{-1}s^{-1}$ [pedestrians per meter per second]) depends on the crowd density (in unit m^{-2} [pedestrians per square-meter]). The use of the fundamental diagram is so elementary in transportation engineering that it is sometimes used as universal truth, without taking the limitations of its validity into consideration:

- *Effect of body sizes:* The average body projection area (onto the floor) scales the density axis and significantly changes the fundamental diagram.
- *Measurement of crowd density:* Is the crowd density measured by counting the number of pedestrians and divided by the area in which the pedestrians were counted, or is the total area divided into small cells, or is the density measured with some other statistical approach?
- *Measurement of flow rate:* Is the specific flow rate measured by counting the pedestrians crossing a line within T seconds and then divided by both T and the length of the line, or is the specific flow rate averaged by adding the speeds of pedestrians within an area and dividing by that area? In any case, what is a suitable value of T ? A small T will result in very scattered data, while a large T on the other hand, will result in a smoothing of the data.
- *Motion characteristics:* Is the motion of the pedestrians mainly into the same direction, or are pedestrians moving into different directions? Are pedestrians moving all the time, or do they sometimes stop without being hindered?
- *Psychological factors:* Is the crowd relaxed or stressed, or even in panic?

See Ref. [26] for details of factors which influences pedestrian route choice. Most of these factors are assumed to influence the fundamental diagram as well.

Even though the fundamental diagram is always presented as a 2-dimensional curve, it is in reality a multi-dimensional relation. To have a fully general fundamental diagram, we need to add (at least) the following quantities:

- Desired speed (not only measured speed).
- Slope of the ground.
- Number of distinct flow directions, and fractions of pedestrians moving in each direction.
- Various psychological and physiological factors [114] (body size, weight, age, gender, calmness, fatigue, etc.).
- Weather conditions (temperature, rain, snow, etc.).
- Environmental conditions (smoky, dark, etc.)

Clearly we need to reduce the system to something with less dimensions. The traditional way to cope with this problem is to assume normal weather conditions, mixed/heterogeneous crowds composed of different gender, age, socioeconomic background, etc., with distributions representative to a larger population, no slope on the ground, and so on. But in some studies, one of the quantities above has been varied and multiple fundamental diagrams have been obtained, e.g.

- for plain ground versus stairs [136],
- for a rolling floor, similar to ship motion [86],
- for normal walking versus crawling [102].

When setting up experiments to obtain fundamental diagrams, it is not possible or practical to add one of the factors above as a third axis of the fundamental diagram. In simulations however, we have more freedom in this regard.

Even though most of the social and physiological factors can be evened out by using a large enough heterogeneous crowd, one has to be careful when transferring the results and using them in a different context from where the results were originally obtained. For example, some factors like body size or desired walking speed vary largely from one country to another [125, 138]. This will result in large differences of the fundamental diagram. Therefore, making flow-density measurements in one country and using

the results for dimensioning the facilities in a different country, could potentially have disastrous consequences.

Another factor that must be taken into consideration is how well the subjects in the crowd are acquainted with each other and with their environment. We illustrate this by two examples:

1. Commuters using the train or metro system in large cities are usually well adapted to high levels of crowding.
2. Huge international gatherings (e.g. sport events or religious gatherings) attended by people from all over the world typically consist of sub-populations with different backgrounds and behaviors.

The level of coordination in the latter example is expected to be far worse than in the first example, and as a consequence, the fundamental diagram and the effective capacity of a pedestrian facility.

4.2 Measurement of Local Densities, Speeds and Flows

The usual definition of density assumes to some extent a homogeneously distributed mass of people, which is approximately true for small areas. For large areas, however, the mass of people can vary from very crowded sub-areas to totally empty sub-areas. This can easily be treated by dividing large areas into a grid of smaller sub-areas, and measure the density in each of these sub-areas. In favor of this pragmatic approach, we chose a more general treatment of large areas, which does not suffer from the boundary effects of using a fixed grid.

The fundamental diagram rests on the assumption that a pedestrians's mobility (speed) depends on the density of pedestrians around him/her. We make the assumption that a pedestrian α is influenced by his/her neighboring pedestrians β with a weight function $f_w(d_{\alpha\beta})$, which decreases with the distance $d_{\alpha\beta}$ between them. When choosing a bell-shaped kernel as a weight function, we end up with the following definition of the *local density*:

$$\rho(\vec{r}, t) = \rho_t^R(\vec{r}) = \frac{1}{\pi R^2} \sum_j \exp[-\|\vec{r}_j(t) - \vec{r}\|^2/R^2]. \quad (4.1)$$

Here, R is a parameter reflecting the size of the area of influence. For practical reasons, one has to scale the local measurements by the factor

$$1/\left[\frac{1}{\pi R^2} \int_{x_0}^{x_1} \int_{y_0}^{y_1} \exp\left(-\frac{(x - r_x)^2 + (y - r_y)^2}{R^2}\right) dx dy\right], \quad (4.2)$$

where $\vec{r} = (r_x, r_y)$ is the central location of the measurement and x_0, x_1, y_0, y_1 are the boundaries of the video-recorded area. For example, in the corners of the recorded area, this factor will approach a value of 4, because 75% of the surrounding pedestrians are outside the video-recorded area.

Note however, even though systematic errors can be compensated for, by such a scaling, there is still a large uncertainty when up to 75% of the data is “invisible”. Therefore, it is recommended to always leave a certain margin around the video-recorded area when determining the local densities, to minimize boundary effects.

The same weighting can be used to define the *local velocity*:

$$\vec{v}(\vec{r}, t) = \vec{v}_t^R(\vec{r}) = \frac{\sum_j \vec{v}_j \exp[-\|\vec{r}_j(t) - \vec{r}\|^2/R^2]}{\sum_j \exp[-\|\vec{r}_j(t) - \vec{r}\|^2/R^2]}. \quad (4.3)$$

Furthermore, we define the *local speed* as

$$v(\vec{r}, t) = v_t(\vec{r}) = \|\vec{v}_t^R(\vec{r})\|, \quad (4.4)$$

and the *local flow* as

$$q(\vec{r}, t) = q_t^R(\vec{r}) = \rho(\vec{r}, t) \vec{v}(\vec{r}, t). \quad (4.5)$$

For $\vec{r} = 0$ and when restricting the area of integration in Eq. (4.2) to $x^2 + y^2 \leq R^2$, we get the result 1/0.63, which implies the following interpretation of R: *The weight of neighboring pedestrians within the radius R is 63%*.

In order to illustrate the relation between the *global density* and the *local density* let us analyze histograms and plot the probability-density functions of the local densities for three different global densities (see Fig. 4.1).

4.2.1 Data Evaluation Method

To gain an intuitive insight into how the local measurements compare to the traditional measurements of the fundamental diagram, hereafter referred to as *global measurements*, we set up computer simulations. For a circular area with 100 m radius, we generate a number of particles corresponding to the global densities $\varrho = 1, 2, 3, \dots, 10 \text{ m}^{-2}$. To avoid boundary effects, measurements are carried out at random locations within 10 m from the center of the circle. Different techniques are used to generate the particles; (a) uniformly distributed (each particle put at a random location without taking into account the positions of the other particles), (b) a regular distribution in a hexagonal grid, and (c) the distribution resulting after a 30-seconds long microsimulation with the social-force model (see Sec. 5.2). Figure 4.2 shows the results from these computer

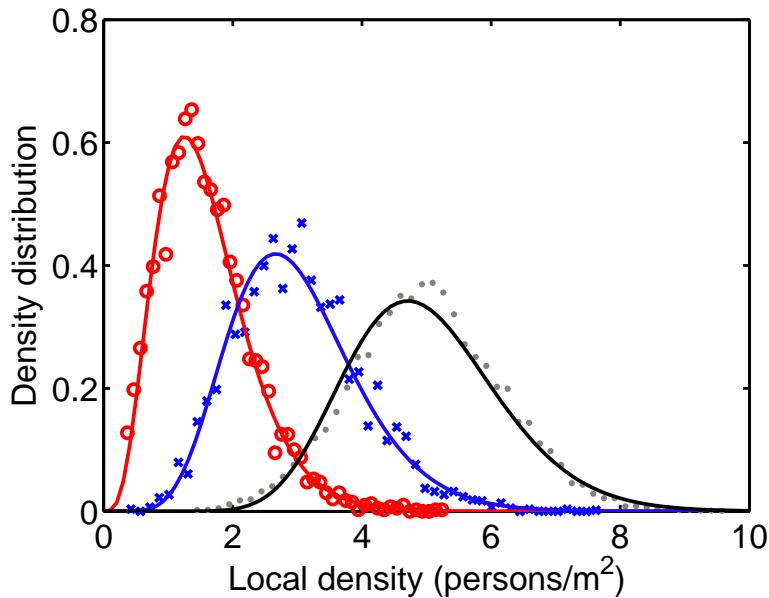


Figure 4.1: Three probability-density functions (PDFs) of the local densities are measured for three different average (global) densities, red: 1.6m^{-2} , blue: 3.0m^{-2} , and black: 5.0m^{-2} . Gamma distributions (solid lines) fit the empirical PDFs well and show the large variation of local densities, which can take values of more than twice the average density.

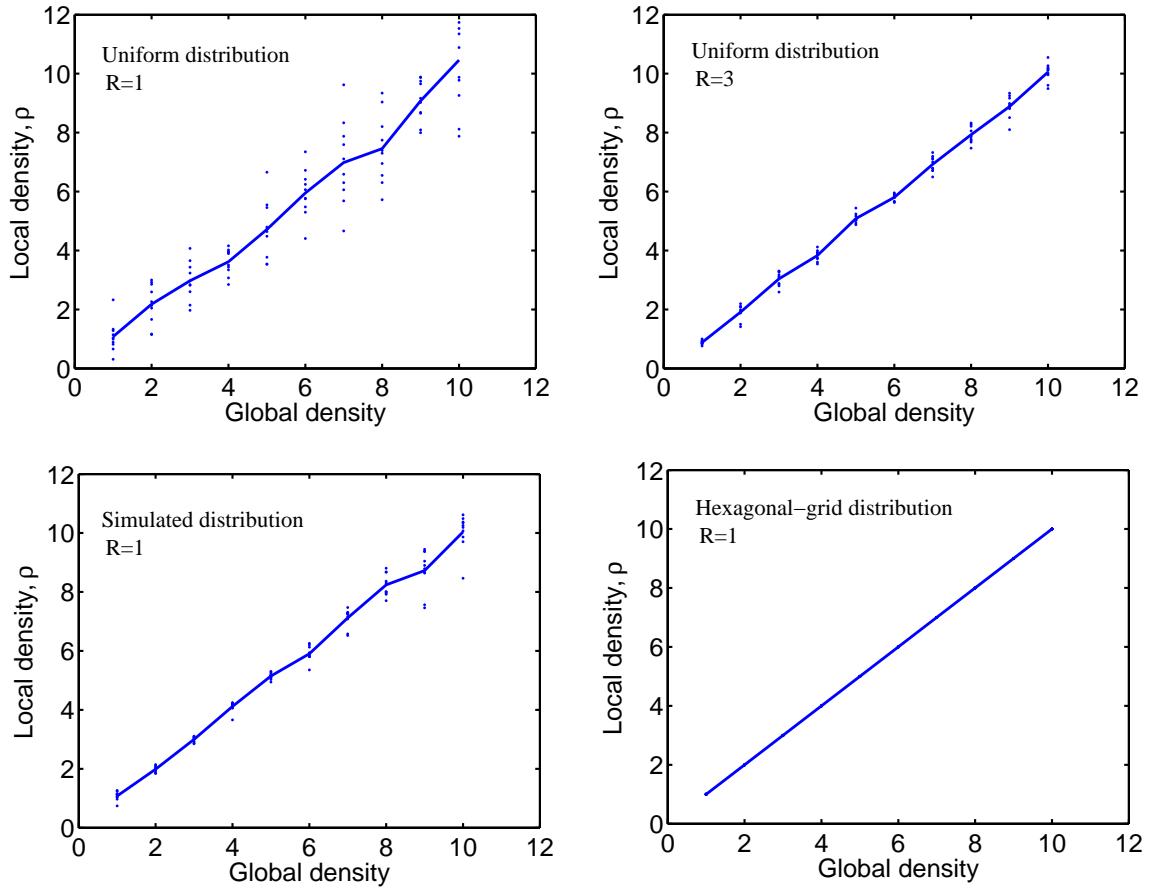


Figure 4.2: In order to illustrate the relation between global and local density, a circle of 100-meter radius is filled with N points according to a given global density $\varrho = N/A = N/(\pi r^2)$. When the circle has been filled with points, 10 random locations are chosen inside the circle, and the local density ρ is measured at these points, and finally plotted as a function of the global density. A line is connected between the average of the 10 local-density measurements and the corresponding global density to show that there is no systematic difference between the two measurement methods. The 10 points are picked within 10-meter radius from the center of the large circle, in order to avoid boundary effects. The way how to distribute the points inside the circle are, Top: uniformly distributed random locations, Bottom right: distributed onto a hexagonal grid, and Bottom left: The result of a simulation with the social-force model.

The largest differences in global vs. local densities are found for the uniform distribution, since there is no minimum distance defined (there may be overlaps).

experiments. For each distribution and for each density value, 10 random locations are used for the measurements (shown as dots). Not very surprisingly, the scattering is lowest for the regular hexagonal-grid distribution and the scattering is highest for the uniform distribution and a small R .

Note that there is no systematic error, since the mean value of the local measurements matches the global measurements. The variance of the local measurements is decreasing with higher values of R . Using $R \geq 3$, the local and global measurements approximately agree. This is not our aim though! We want to grasp the “true” density, i.e. only the neighboring pedestrians β , around a pedestrian α , that are really influencing the mobility of pedestrian α . Therefore, a suitable R should be $R < 3$. What is the best value of R will be elaborated later on.

4.3 Empirical Findings

Now, we will start applying our theory to real videos. We use the videos from the Muslim pilgrimage to Makkah, in January 2006, explained more in detail in Sec. 7. The videos are analyzed with the methods described in Sec. 2.3, and the result is a large dataset. 45 minutes (1.3 GB) of video recordings are analyzed, resulting in a dataset, containing all positions $\vec{r}_\alpha(t)$ and velocities $\vec{v}_\alpha(t)$, of all pedestrians α , in the frames at the times $t = n\Delta t$. The framerate of 8 frames per second gives the timestep $\Delta t = 1/8$ s.

4.3.1 Dealing with Umbrellas

The time interval evaluated in our analysis starts with low density and smooth flow, which increases all the way to densities high enough to trigger a crowd disaster (see Sec. 7.4). Therefore, we have a large span of different conditions during the 45 minutes of evaluated data. However, there is one peculiarity in the videos, which was not dealt with in Chapter 2, namely people carrying umbrellas, for protection against sunshine. Unfortunately, the time period with the maximum crowd density approximately agrees to the time with most intensive sunshine. Therefore, we can not ignore this problem, but rather need to find a way to deal with it. The average umbrella diameter is measured to be 1.0 m, and deviations from this size are rarely seen in the videos. Each umbrella covers an unknown number of pedestrians.

To compensate for umbrellas, the following approach is used: Assume that the total area in the video recording is A [m^2], the area per person is A_1 [m^2] and the area of an umbrella is $A_2 = \pi 0.5^2$ m^2 . This gives, on average, $k = A_2/A_1$ covered persons under

each umbrella. If n is the number of visible persons, and m the overall number of invisible persons per umbrella, $N = n + km$ corresponds to the total number of persons. Suppose that a fraction p of the pilgrims are carrying umbrellas, which gives $m = pN$. From the video analysis, we estimate the global density to be $\varrho' = n/A$. Putting everything together, we have,

$$N = n + km = A\varrho' + \frac{N^2 p A_2}{A}. \quad (4.6)$$

Solving for the actual crowd density $\varrho = N/A$ we get

$$\varrho = N/A = \frac{1 - \sqrt{1 - 4pA_2\varrho'}}{2pA_2}. \quad (4.7)$$

Note that, since N results from a second-order equation, we have two roots. However, only the negative root gives reasonable values, therefore we omit the positive root.

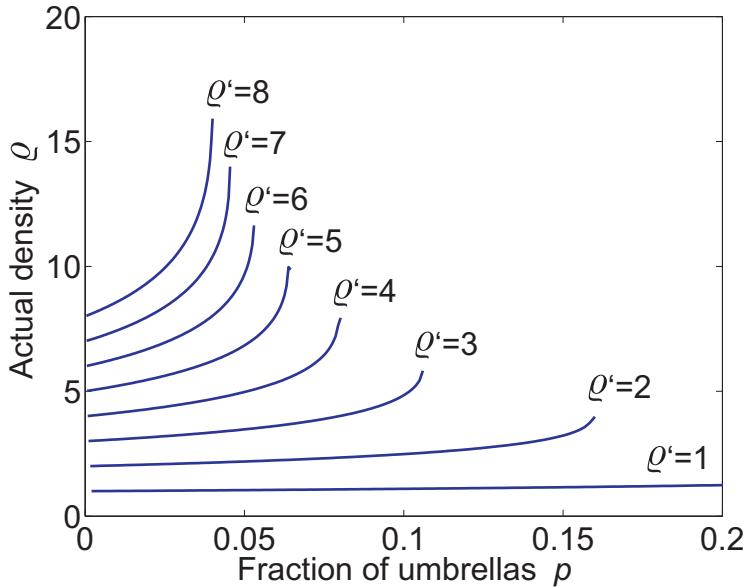


Figure 4.3: Actual (real) density ϱ , as a fraction p of umbrellas, plotted for different values of measured density ϱ' .

Equation (4.7) is applied to different values of measured density ϱ' , to see how the real density ϱ scales with p (see Fig. 4.3).

The theory outlined above works well if the parameter values are known. The average area of an umbrella can easily be determined from the videos, but the fraction of persons holding umbrellas, p is not so easy to obtain. Therefore, we will now present an

alternative theory to handle the problem with umbrellas. It works under the condition that we do not need to know the average density, but rather we want to estimate which areas contain umbrellas and which areas are free from umbrellas. With the same logic as applied in Sec. 4.2 and Eq. (4.2), locations close to umbrellas will produce too low local-density estimates. Therefore, if we

1. pick a number of random locations, within the video-recorded area,
2. calculate the local densities ρ_i for all location i ,
3. sort the values by decreasing density, and
4. filter out the fraction γ of the lowest density values,

the remaining locations are likely to have no nearby umbrellas.

To test this hypothesis, we again starts a batch of microscopic simulations with the social-force model, assuming different global densities. Filtering out half of the values ($\gamma = 50\%$), there is still a significant underestimation of the local-density measurements compared to the actual densities. With γ as high as 95%, the influence of umbrellas is successfully eliminated. Therefore, from now on, we will always use $\gamma = 95\%$ to filter out umbrellas, in order to get reliable results.

Let us now check the plausibility of the methods proposed above to correct for umbrellas in real data. In order to investigate how the speed-density and flow-density diagrams depend on the cutoff value γ , we have evaluated them for different values of γ (see Fig. 4.5). We find that, while the speed-density relationships vary relatively little with a variation of γ , there is a significant increase of the local flows for greater γ values, particularly in the intermediate density range. This indicates that fitting speed-density data may lead to unreliable conclusions regarding the flow-density relationship, while fitting flow-density data would lead to good velocity-density fits as well. The reason is that the density enters twice and in a multiplicative manner into the flow (as product of density and speed).

Moreover, we have determined the globally averaged speed and the global flow as a function of the global density for different assumed fractions of umbrellas p (see Fig. 4.6).

Let us now address the question, how the *actual* fraction of umbrellas can be determined from the video recordings. For this, we use the fact that the relationship between the local speed and the local density should not be significantly changed by the fraction p of umbrellas. We, therefore, define a reference relationship $\rho_{\text{ref}}(v)$, measure the velocity

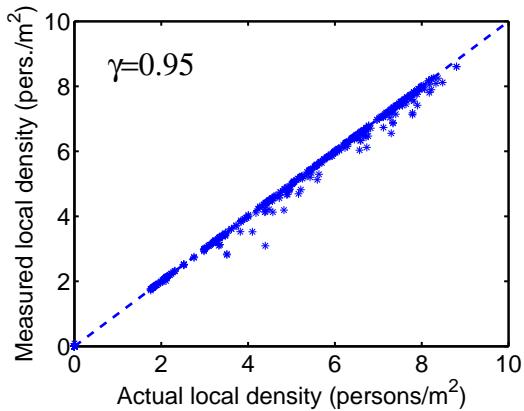
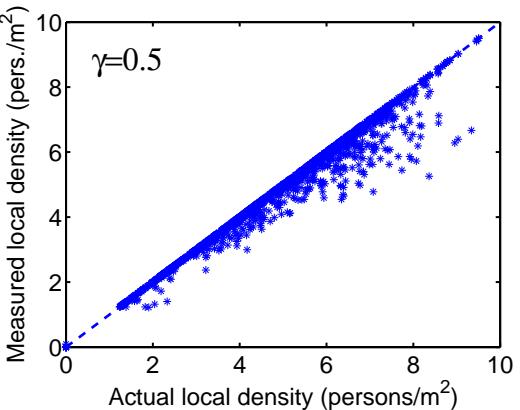
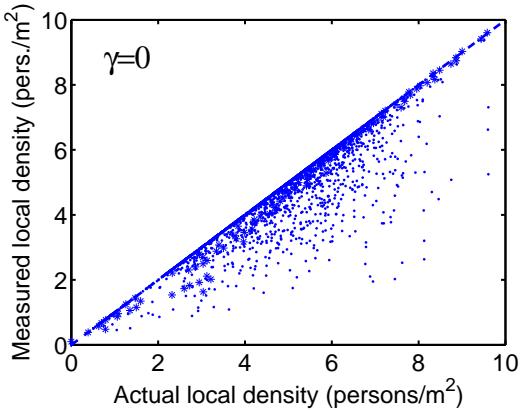


Figure 4.4: To determine the influence of umbrellas on the measurements, simulations are carried out with the social-force model, assuming that 2% of the pedestrians carry an umbrella with a radius of 50 cm. First, the actual local density is measured for all pedestrians. Then, the estimated (measured) local density is determined by removing all pedestrians within a 50 cm radius of any umbrella. In an attempt to filter out the influence of umbrellas, a fraction γ of the lowest measured local densities are filtered out. Finally, the measured local density is plotted against the actual (true) local density, for $\gamma = 0$, 0.5, and 0.95. It turns out that only $\gamma = 0.95$ gives an acceptable accuracy.

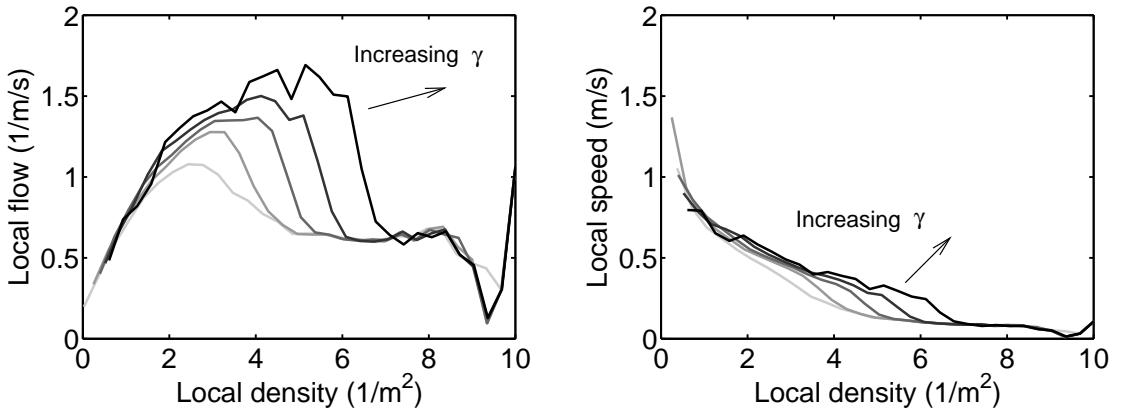


Figure 4.5: Empirical relationships between the local speed and the local density (left) and between the local flow and the local density (right) for various values $\gamma \in \{0\%, 25\%, 50\%, 75\%, 95\%\}$. The cutoff value γ is supposed to correct for the influence of umbrellas. It is clear that the flow values drop, if pedestrians covered by umbrellas are not suitably accounted for, which is the case for small values of γ .

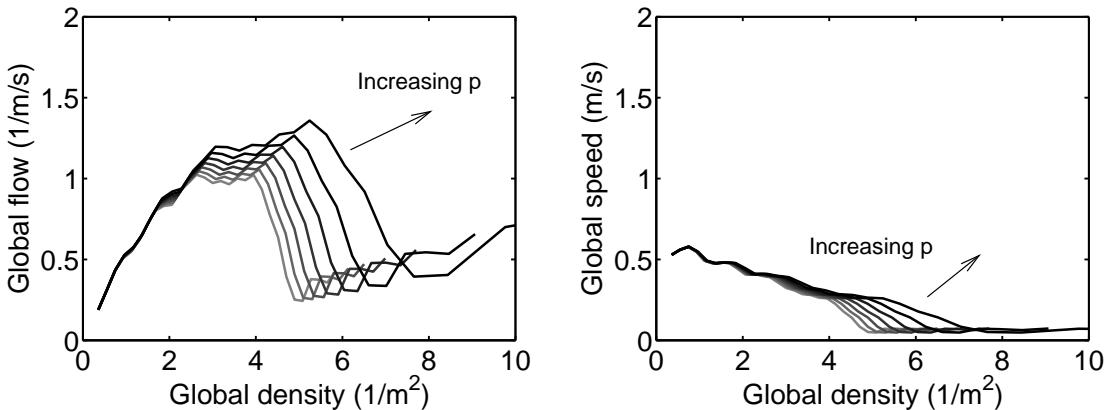


Figure 4.6: Empirical relationships between the globally averaged speed and the global density (left) and between the globally averaged flow and the global density (right) for various values of $p \in \{0\%, 1\%, 2\%, \dots, 6\%\}$, using formula (4.7). As a larger assumed fraction p of umbrellas implies a larger number of hidden pedestrians, it is clear that the flow must increase with the value of p .

v and density ρ , and calculate the fraction p of umbrellas that leads to consistent density values. In detail, the procedure is as follows:

- As a reference relationship $\rho_{\text{ref}}(v)$, we use a fit curve (obtained from empirical data) of the local density as a function of the local speed, determined for $R = 1$ and $\gamma = 95\%$. According to our previous numerical studies, the large cutoff value γ should eliminate the influence of umbrellas well.
- At a given time t , we pick 1000 random locations \vec{r}_i .
- For these locations, we determine the local densities $\rho'|_{\gamma=0}(\vec{r}_i)$ and the local speeds $v|_{\gamma=0}(\vec{r}_i)$ with $R = 1$ m and $\gamma = 0$. The value $\gamma = 0$ ignores umbrellas and results in density values $\rho' \leq \rho$, but while the measured speeds should be correct.
- We then estimate the corrected densities as $\rho(\vec{r}_i) = \rho_{\text{ref}}(v(\vec{r}_i))$.
- Next, we determine the global density $\varrho(t)$ as average of the corrected local densities $\rho(\vec{r}_i)$, i.e. $\varrho(t) = \langle \rho(\vec{r}_i) \rangle$. Similarly, we obtain the global density $\varrho'(t)$ ignoring umbrellas as average of the densities $\rho'|_{\gamma=0}(\vec{r}_i)$, i.e. $\varrho'(t) = \langle \rho'|_{\gamma=0}(\vec{r}_i) \rangle$.
- Finally, we can estimate the fraction of umbrellas via the formula

$$p(t) = \frac{\varrho(t) - \varrho'(t)}{\varrho^2(t)A_2}, \quad (4.8)$$

which follows from Eq. (4.7). As the radius of most umbrellas is 0.5 m, we use the value $A_2 = \pi 0.5^2 \text{ m}^2$.

The empirically determined fraction $p(t)$ of umbrellas as function of time t is shown in Fig. 4.7. It turns out that the fraction of umbrellas increases after 10:30am and reaches its maximum at noon time, when the impact of the sun is strongest. This shows the plausibility of our procedure, which has been checked by manual counts.

4.3.2 Relationships between Densities, Velocities and Flows

The question of choosing a suitable value of R remains. We plot the fundamental diagram for three different definitions of R (see Fig. 4.8). Larger values of R will smoothen the data, and since we are particularly interested in high densities we will use the value $R = 1$, which is sensitive enough to determine density differences of large average densities.

By using $R = 1$, we obtain the fundamental diagram and compare it to the fundamental diagrams of other authors (Fig. 4.9).

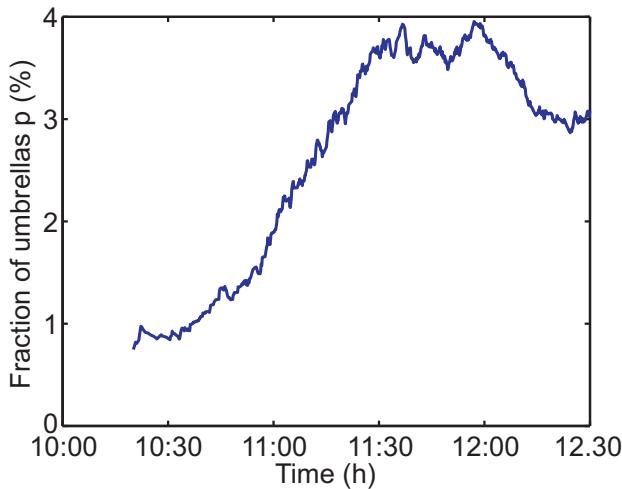


Figure 4.7: Fraction p of pilgrims carrying umbrellas as a function of time. One can see that p is four times larger around noon time, compared to an hour earlier.

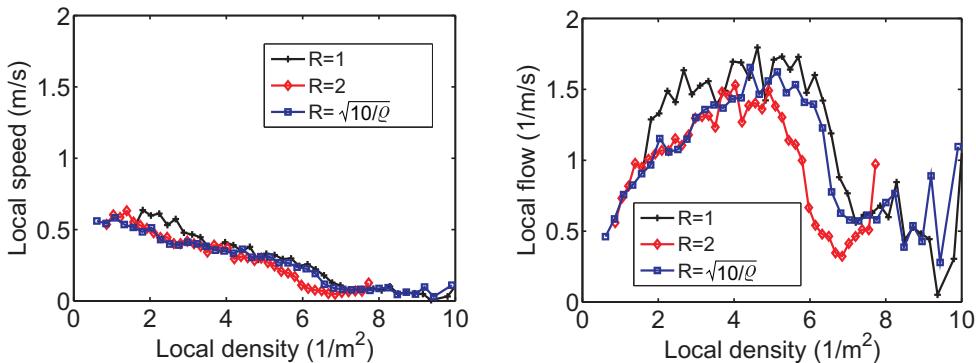


Figure 4.8: Average local speed (left) and flow (right) as a function of the local density ρ . Curves are plotted for constant $R = 1 \text{ m}$, $R = 2 \text{ m}$, and for the density-dependent specification $R = \sqrt{10/\rho}$, corresponding to a constant expected number of pedestrians in the area $A_R = \pi R^2$. One can see that smaller values of R tend to imply larger flow values, as the averaging is performed over smaller areas so that extreme values are not averaged out. In other words: The variation of local densities is smaller the larger the value of R . As cutoff value we have used $\gamma = 0.95$.

Note that our curve looks very different from most of the other curves. The reason is manifold:

- Our data and the data from Mori and Tsukaguchi [100] show higher flows at higher densities than the other data. This can be explained by smaller body sizes and smaller individual space requirements in Asian crowds [126], compared to Western crowds.
- The motivation to move forward is very strong in our data, since for most pilgrims it is a once-in-a-lifetime event. In contrast, the other data were determined for everyday situations.
- Our dataset is unique as it contains data in the very-high-density regime.
- Our dataset is measured in a 20 m wide area, which allows for large fluctuations and large-scale patterns (see Secs. 3.6 and 3.7), while the other data are measured at much smaller cross-sections.

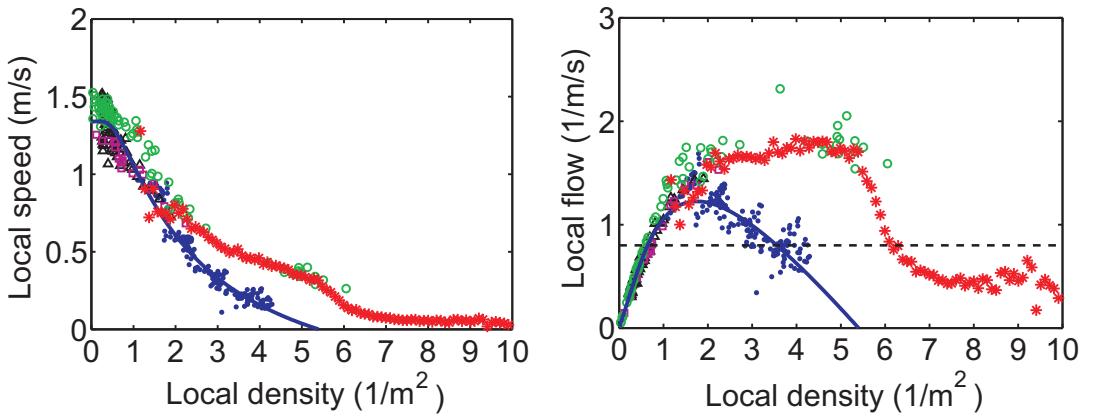


Figure 4.9: The local flow-density relationship as compared to flow-density data of other studies. Local speed (left) and local flow (right) as a function of local density. Symbols correspond to empirical data by Fruin [36] (black triangles), Mori and Tsukaguchi [100] (green circles), Polus *et al.* [112] (pink squares), and Seyfried *et al.* [117] (blue dots). The blue solid fit curve is from Weidmann [136]. Our data are displayed as red stars.

4.4 Fundamental-Diagram Model

Up to now, various aspects of the fundamental diagram have been discussed, but we have focused on the macroscopic quantities (density, flow, velocity) neglecting that the fundamental diagram reflects the aggregated behaviors of single pedestrians. Therefore, let us now make an attempt to explain the fundamental diagram with a bottom-up approach starting with pedestrian interactions.

To connect the aggregated crowd density to a local scale, let us approximate the mean distance between the center-of-masses of a pedestrian α and the closest pedestrian β by $d = \langle d_{\alpha\beta} \rangle = 1/\sqrt{\varrho}$, where ϱ is the global (average) crowd density. Note that this would hold only if the pedestrians were distributed into a square lattice, but for other density distributions it will serve as a fair approximation. The net distance is defined as $\hat{d} = d - 2r$, where $r = 1/(2\sqrt{\varrho_{max}})$ is the effective radius of a pedestrian, and ϱ_{max} is the largest measured density. For this reason, we include only studies [54, 100, 44, 128, 9, 136] containing a complete range of crowd densities.

Assuming that the predecessor β would suddenly stop, it would take $\hat{T} = \hat{d}/v_\alpha$ seconds before a physical encounter with pedestrian α occurs, if v_α is the speed of pedestrian α . We now show how the net-time headway \hat{T} depends on the global crowd density ϱ by applying the above scheme to empirical data determined from different authors (see Fig. 4.10). Note that \hat{T} most often saturates at a constant value. However, in our data [54] there is a transition at very high crowd densities, where \hat{T} suddenly *increases*. This can be interpreted in at least two ways:

- *Hypothesis 1:* When the crowd density is very high, pedestrians start to have fear of crushing or asphyxia [143], and therefore want to increase the space around themselves (leading to higher net-time headways \hat{T}).
- *Hypothesis 2:* If the space in front of a pedestrian is too small (or the velocity is too low) it will no longer be possible to make normal steps. Rather, pedestrians will completely stop until they have gained enough space to make a step.

In previous work [143], *Hypothesis 1* has been used. In this study, however, we will investigate *Hypothesis 2*. This interpretation would naturally explain the empirically observed stop-and-go waves (see Sec. 3.6), and would further imply: *above 5 persons per m², the fundamental diagram will no longer describe the dynamics of the crowd well, since the flow rate is then alternating between movement and standstill rather than continuous.*

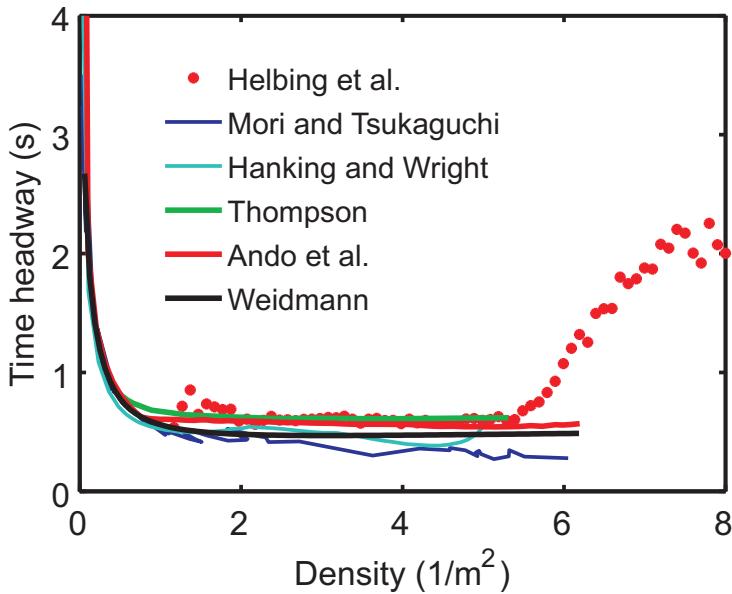


Figure 4.10: Net-time headways \hat{T} as a function of crowd density ϱ for a number of empirical studies. Note that \hat{T} most often saturates at a constant value. In our data [54], however, there is a transition for high crowd densities where \hat{T} suddenly increases. The net-time headway is calculated as $\hat{T}(\varrho) = \left(1/\sqrt{\varrho} - 1/\sqrt{\varrho_{max}}\right)/v(\varrho)$.

In an attempt to unify all fundamental diagrams in the same framework, the following scheme is proposed:

Each pedestrian α has a free speed $v_\alpha^0 = v_{max}$ (which is an upper speed limit, occurring when $\rho \rightarrow 0$). Each pedestrian also has a lower limit v_{min} of the speed. For $v < v_{min}$, pedestrians can no longer make normal steps, and would rather stop completely. For simplicity, these bounds are supposed to be the same for all pedestrians. It has been reported in Ref. [66] that the free (unconstrained) headways are exponentially distributed, where the constrained headways on the other hand, are limited by a desired minimum headway. Therefore, we propose that the net-time headway \hat{T} is the key control parameter for the fundamental diagram. That is, pedestrians will decrease their speeds, if necessary, to assure a constant lower limit of the net-time headway \hat{T} . The

fundamental diagram can now be specified as:

$$v(\varrho) = \frac{d - 2r}{\hat{T}} = \frac{1/\sqrt{\varrho} - 1/\sqrt{\varrho_{max}}}{\hat{T}} \quad (4.9)$$

and bounded by $[v_{min}, v^0]$.

We know from Sec. 4.2 that each average density ϱ corresponds to a distribution of local densities ρ , and we therefore approximate the density distribution with a Gaussian distribution $\mathcal{N}(\varrho, \sqrt{\varrho/3})$ with mean ϱ and standard deviation $\sqrt{\varrho/3}$ (which has been obtained from the data presented in Fig. 4.1).

The advantage of this approach is that fundamental diagrams can be generated without using any arbitrary fit function. Figure 4.11 shows generated fundamental diagrams from Eq. (4.9) with the parameters $\hat{T} = 0.5$ s, $\varrho_{max} = 5.4$ m $^{-2}$, and for different values of the free speed $v^0 = v_{max}$. Since v^0 only gives the upper limit of the velocity, fundamental diagrams with different v^0 converge at high enough crowd densities, when the movement has been transformed from individual walking to walking which is constrained by other pedestrians.

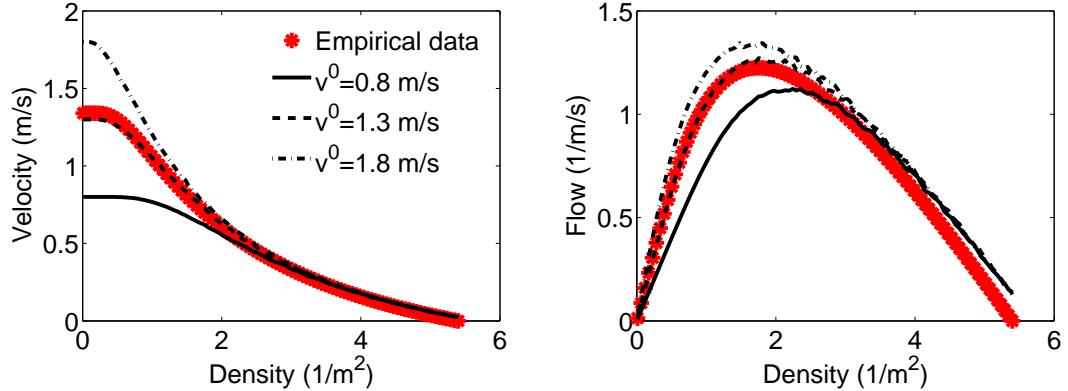


Figure 4.11: Fundamental diagrams generated by Eq. (4.9), for different free speeds v^0 .

Note how they all converge for large densities. As a comparison, the empirical fit curve by Weidmann [136] is shown.

According to *Hypothesis 2*, defined above, we get an extra constraint, saying that pedestrians will stop walking if they are too close to other pedestrians, which happens for $\rho \geq \varrho_{max}$ (physical interaction). They will then resume walking again when they have enough space L for taking a step. Since one step needs approximately $L = 0.5$ m, we get a new net-time headway $\hat{T}' = L/v_{min} \approx 20\hat{T}$ s, whenever $\rho \geq \varrho_{max}$ (see Fig. 4.12).

The fraction of pedestrians which are physically colliding with others, can be measured by integrating the probability-density-function (see Fig. 4.12 (left)) of the Gaussian distribution

$$f_{stop} = \int_{\varrho_{max}}^{\infty} \mathcal{N}(\varrho, \sqrt{\varrho/3}) d\varrho. \quad (4.10)$$

Then, the mean net-time headway (see Fig. 4.12 (right)) is given by the fraction of stopped pedestrians as

$$\langle \hat{T} \rangle = (1 - f_{stop})\hat{T} + f_{stop}20\hat{T}. \quad (4.11)$$

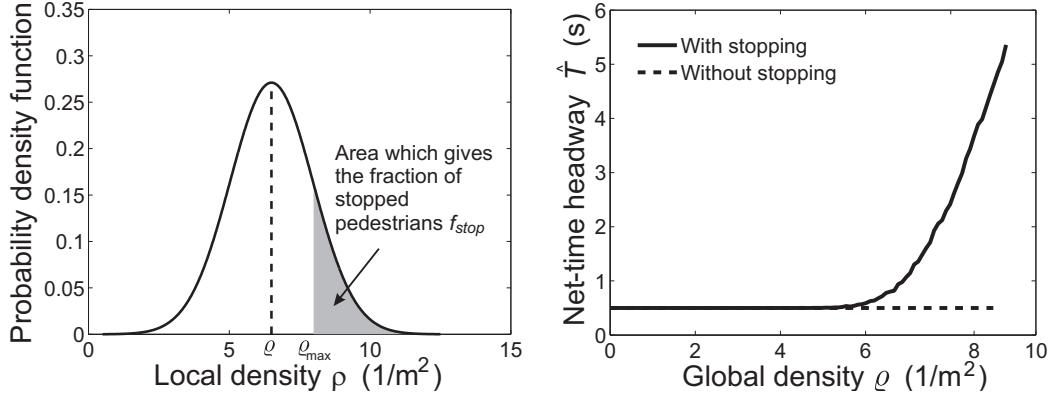


Figure 4.12: Left: The mean net-time headway $\langle \hat{T} \rangle$ is obtained via the fraction of pedestrians who are physically colliding with others, and is therefore stopping and temporarily increasing their net-time headway. This fraction is obtained by integrating over the probability-density-function of the local-density distribution, starting at local densities ρ that are higher than the maximum global density ϱ . Right: The mean net-time-headway $\langle \hat{T} \rangle$ (solid line) as a function of the global density ϱ . The net-time headway without stopping is displayed as a dashed line.

We now apply the method outlined above on different empirical fundamental diagrams. In all cases we use $\hat{T} = 0.5$ s and $v_{min} = 0.06$ m/s. Starting with Weidmann's [136] fundamental diagram, we have the parameters $\varrho_{max} = 5.4$ m^{-2} and $v^0 = 1.34$ m/s, which is displayed in Fig. 4.13 (top) together with our curve, obtained by Eq. (4.9).

Next, we apply our method on the fundamental diagrams from Refs. [54] and [100] and obtain the results presented in Fig. 4.13 (middle and bottom).

To conclude, the scale of the fundamental diagram depends on the maximum density ϱ_{max} and, therefore, on the body sizes.

Furthermore, the shape of the fundamental diagram depends on the density variations. The capacity drop for high densities is sharper, the higher the variation of the densities. All three fundamental diagrams [136, 100, 54] share the same net-time headway $\hat{T} = 0.5$ s.

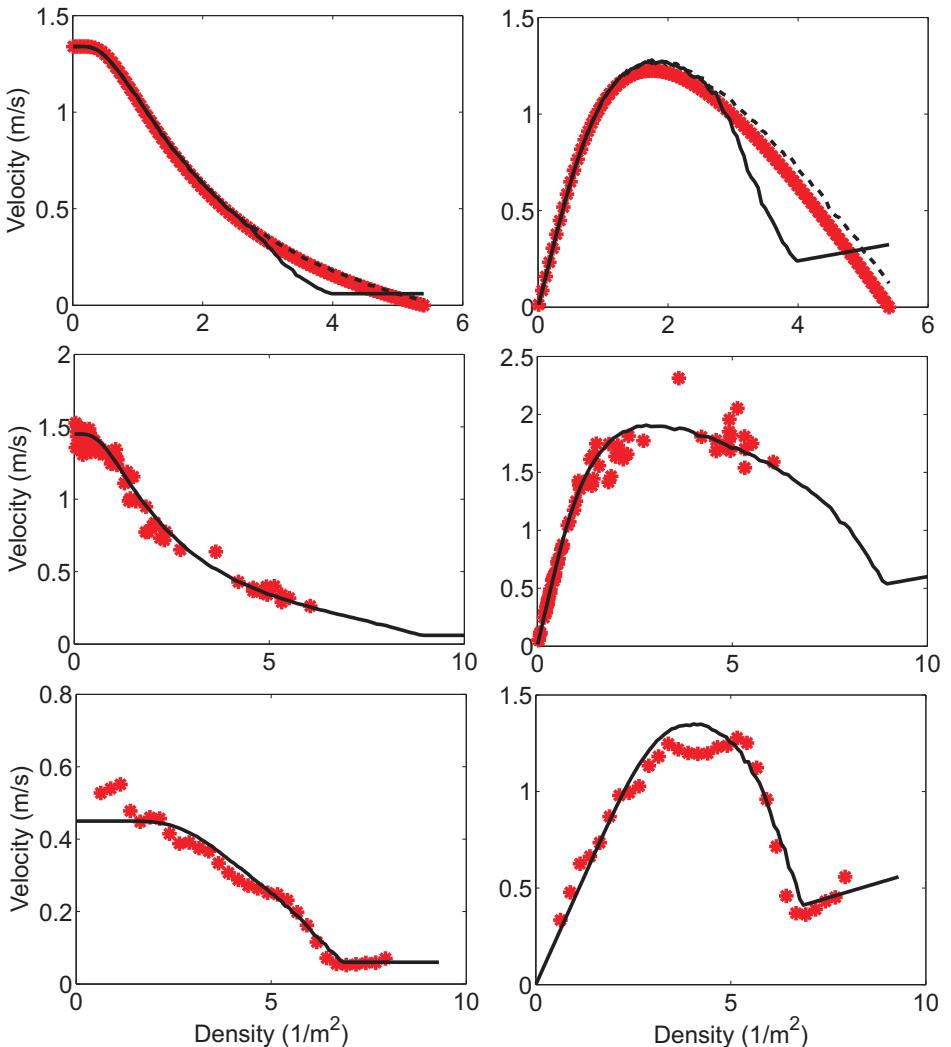


Figure 4.13: Diagrams of velocity (left) and flow (right) vs. density, generated by Eq. (4.9), assuming a constant net-time headway $\hat{T} = 0.5$ s and a minimum velocity $v_{min} = 0.06$ m/s. Markers represent the empirical data and solid lines represent our generated fit functions. Top: Data from Weidmann [136] (red markers) as compared to a fit function generated with the parameters $\varrho_{max} = 5.4 \text{ m}^{-2}$ and $v^0 = 1.34 \text{ m/s}$. The dashed line shows the result when it is assumed that no pedestrians are stopping, i.e. $f_{stop} = 0$. Middle: Data by Mori and Tsukaguchi [100] compared to a fit function generated with the parameters $\varrho_{max} = 12 \text{ m}^{-2}$ and $v^0 = 1.45 \text{ m/s}$. Bottom: Data from Ref. [54], compared to a fit function generated with the parameters $\varrho_{max} = 9.3 \text{ m}^{-2}$ and $v^0 = 0.45 \text{ m/s}$. Each velocity $v(\rho)$ is bounded by $[v_{min}, v^0]$ and each global density ϱ corresponds to a distribution of local densities ρ . Therefore, the global velocity $V(\varrho) = \langle v(\rho) \rangle$ becomes smoothed.

5 Modeling and Simulation

5.1 Introduction

In the previous chapters we have focused mainly on (extraction and analysis of) empirical data of pedestrian motion. We will now address the question of how we can create mathematical models and run computer simulations of pedestrians. In the past, pedestrian simulations have mainly been used to *qualitatively* reproduce and understand various aspects of crowds. Nowadays however, neither the computing performance nor the amount and detail of available empirical data restrict us from aiming at reproducing crowd dynamics *quantitatively* as well.

Starting with a general introduction to pedestrian modeling, we will then introduce the social-force model and show how this model can be calibrated and evaluated with empirical data that have been obtained with the methods from Chapter 2.

Modeling of pedestrian motion is performed for different reasons: to quantitatively investigate evacuation dynamics, to assess public facilities, or simply to gain more knowledge about how pedestrians interact with each other and how the design of the infrastructure affects the flow of pedestrians.

Among the first ideas to simulate interacting agents in a swarm-like way was proposed by Reynolds [115] with his *Boids* model.

To model pedestrians properly, several levels of abstraction need to be taken into consideration, reaching from the lowest (operational) level of crowd dynamics, to higher levels including route choice [26, 64, 22] and decision making [145, 65, 11]. There are also various models that aim at quantifying the space in a hierarchical way, for example *space syntax* [108] and the *chronotopic model* [15].

We will now focus on the lowest level of pedestrian modeling, i.e. crowd dynamics. Simplifying matters, the higher levels are based on conscious reasoning, while the lowest level is mainly based on subconscious actions and reactions. This makes it possible to create simple stimulus-response models describing how a pedestrian will automatically respond to different impulses and to the environment. Such models can mimic how

pedestrians interact with each other and with the environment. Surprisingly, such simple models are able to quite realistically reproduce the complex dynamics of crowds [53, 51].

Another interesting behavioral aspect underlying crowd dynamics is how people learn and optimize walking and interaction patterns. This aspect of crowd dynamics is addressed in Refs. [21, 59]

In the past years, a considerable variety of approaches to model pedestrian movement have been proposed.

Here, a short overview will be given. For more comprehensive reviews, the reader is encouraged to read Refs. [26, 59, 51, 52]. The different approaches to pedestrian modeling can be classified according to their level of abstraction:

- *Microscopic* models describe each pedestrian as a unique entity with its own properties.
- *Macroscopic* models delineate the average or aggregate pedestrian dynamics by densities, flows, and velocities as functions of space and time.
- *Mesoscopic* (gas-kinetic) models are in between the two previously mentioned levels, taking into account the velocity distribution.

Further, models can also be classified by their respective detail of description:

- *Discrete space* models sub-divide the environment into a lattice, and the spatial resolution of the model is limited by the cell size of the lattice.
- *Continuous space* models describe the spatial resolution down to an arbitrary level of detail.

Also the time in the model can be either discrete or continuous. Let us now briefly introduce a number of popular pedestrian models.

5.1.1 Agent-Based Models

A class of models which is popular in the computer-science community is agent-based models [104, 101]. These models are characterized by a high level of autonomy of the simulated pedestrians, where each pedestrian is controlled by a set of rules. The advantages with these kind of models are that the motion can look very realistic and that the agents can be adaptive and possess a high degree of artificial intelligence. This also make agent-based models suitable for crowd animation [133, 113]. A disadvantage is

that these kind of models tend to be very complicated, which makes it hard to approach them analytically, and they typically also need a lot of computational effort.

5.1.2 Social-Force Model

The social-force model [46, 57, 53, 51] is a microscopic model, which is continuous both in space and time. It is influenced by Newtonian mechanics, generalized to the motion of pedestrians.

The forces consist of repulsive forces with respect to other pedestrians and boundaries, friction forces, attractive forces among group members, and driving forces related to desired velocities. A superposition of all these forces gives a resultant force which determines the acceleration of the pedestrians. The full details of the model are given in Sec. 5.2.

5.1.3 Cellular Automata Models

Another popular approach to pedestrian modeling is based on cellular automata [21, 20, 95, 88], which is a microscopic model, discrete both in time and space. The exact specification of these models differs, but the common idea is to divide the walkable space into a lattice, where each cell has an area corresponding to the size of a human body, approximately $40\text{ cm} \times 40\text{ cm}$. Each cell can either be occupied by *nobody* or by *one* pedestrian. The movements of pedestrians are performed by iterating the time in steps of about 0.3 s. In each time step the pedestrians can move to unoccupied neighboring cells. However, even though the basic idea of the cellular automata models is simple, it often becomes complex with many rules for how the movement should be performed.

Since the cellular automata models are discrete both in time and space, and due to the fact that they use only local interactions, they are often used for simulating large crowds.

One drawback of the cellular automata models is that the accuracy of the dynamics is questionable, especially for high crowd densities, because the pedestrian bodies are treated as incompressible, and phenomena like clogging due to high pressures or turbulent high density flow are not well reproduced, if at all.

5.1.4 Fluid-Dynamic Models

When the crowd density is high, flows of pedestrians remind of fluids flows. Therefore, a macroscopic approach to crowd modeling is to use fluid-dynamic models [47, 72, 132]

adapted to the simulation of pedestrians. An advantage of fluid-dynamic modeling of pedestrians is that it becomes possible to make analytical evaluations of changes in the infrastructure or changes of boundary conditions.

5.1.5 Queueing Models

Queueing models [135, 94] make further simplifications of crowds. They are used to analyze how pedestrians are moving around in a network of modules, where the nodes and links can, for example, be doors and rooms, or intersections and roads. It is important to stress that the dynamics inside each node is not explicitly taken into consideration.

The idea is rather to grasp how the different modules are interacting with each other, by analyzing queues in the system. Each node has a certain “service rate” and pedestrians move to the next queue as soon as they have been “served”.

5.2 Social-Force-Model Specification

Compared to mechanics and physics, the “particle interactions” of pedestrians have to consider the additional aspects of free will and reasoning. Our approach is to separate pedestrian motion into two categories,

- **Conscious:** This category includes reasoning and planning.
- **Subconscious:** This category includes automatic responses to stimuli from the environment and from other pedestrians.

We are focussing mainly on the *subconscious* part, and replace the whole *conscious* part by one simple objective: The assumption that each pedestrian wants to reach a certain pre-defined location. Therefore, re-planning, multi-objectives, adaptivity to the environment, and memory effects are omitted by now. It has been shown that by using this simplification, even complex self-organization phenomena such as herding, lane formation and oscillations of walking directions at narrow bottlenecks can be reproduced. Note however, that the reasoning part will be addressed in more detail in future work, and is by no means regarded as superfluous.

The social-force model relates to an idea by Lewin [92] that human behaviors are driven by *social fields* or *social forces* and was put into mathematical equations by Helbing [46]. It has been applied to opinion formation [48], pedestrian motion [46], and vehicle traffic [130].

The social-force model for pedestrians assumes that each pedestrian α is trying to move in a desired direction \vec{e}_α with a desired speed v_α^0 , and that he/she adapts the actual velocity \vec{v}_α to the desired velocity, $\vec{v}_\alpha^0 = v_\alpha^0 \vec{e}_\alpha$ within a certain relaxation time τ_α . The velocity $v_\alpha(t) = d\vec{r}_\alpha/dt$, i.e. the temporal change of the location $\vec{r}_\alpha(t)$, is itself assumed to change according to the acceleration equation

$$\frac{d\vec{v}_\alpha(t)}{dt} = \vec{f}_\alpha(t) + \vec{\xi}_\alpha(t), \quad (5.1)$$

where $\vec{\xi}_\alpha(t)$ is a fluctuation term and $\vec{f}_\alpha(t)$ the systematic part of the acceleration force of pedestrian α , given by

$$\vec{f}_\alpha(t) = \frac{1}{\tau_\alpha}(v_\alpha^0 \vec{e}_\alpha - \vec{v}_\alpha) + \sum_{\beta(\neq\alpha)} \vec{f}_{\alpha\beta}(t) + \sum_i \vec{f}_{\alpha i}(t). \quad (5.2)$$

The terms $\vec{f}_{\alpha\beta}(t)$ and $\vec{f}_{\alpha i}(t)$ denote the repulsive forces describing the attempts to keep a certain safety distance to other pedestrians β and obstacles i . The fluctuation term $\vec{\xi}_\alpha(t)$ reflects random behavioral variations arising from deliberate or accidental deviations from the average way of motion. The above equations are nonlinearly coupled Langevin equations. These equations can be solved with a computer by various numerical solvers, ranging from the simple 1st order Euler's method to more complicated iterative methods like the Runge-Kutta method. In very crowded situations, additional physical contact forces come into play [53].

For the time being, we will assume a simplified interaction force, consisting of an angular-dependent part and a distance-dependent part (see Fig. 5.1) of the form

$$\vec{f}_{\alpha\beta}(t) = w(\varphi_{\alpha\beta}(t)) \vec{g}(d_{\alpha\beta}(t)), \quad (5.3)$$

where $\vec{d}_{\alpha\beta} = \vec{r}_\alpha - \vec{r}_\beta$ is the distance vector pointing from pedestrian β to pedestrian α and $\varphi_{\alpha\beta}$ the angle between the normalized distance vector $\hat{\vec{d}}_{\alpha\beta} = \vec{d}_{\alpha\beta}/\|\vec{d}_{\alpha\beta}\|$ and the direction $\hat{\vec{v}}_\alpha = \vec{v}_\alpha/\|\vec{v}_\alpha\|$ of motion of pedestrian α , i.e. $\cos(\varphi_{\alpha\beta}) = \hat{\vec{v}}_\alpha \cdot \hat{\vec{d}}_{\alpha\beta}$. It has been suggested that the reaction of pedestrians to what happens in front of them is much stronger than to what happens behind them. This can be reflected by the function

$$w(\varphi_{\alpha\beta}(t)) = \left(\lambda_\alpha + (1 - \lambda_\alpha) \frac{1 - \cos(\varphi_{\alpha\beta})}{2} \right). \quad (5.4)$$

Here, λ_α with $0 \leq \lambda_\alpha \leq 1$ is a parameter which grows with the strength of interactions from behind.

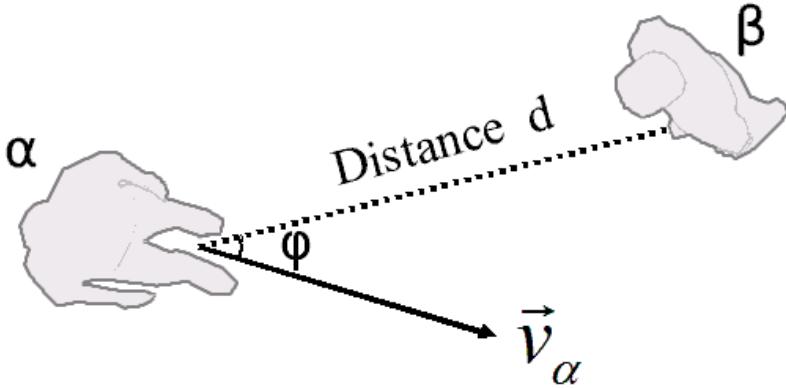


Figure 5.1: A simplified interaction force acting on pedestrian α with respect to pedestrian β is assumed. It has one angular-dependent part and one distance-dependent part.

Later on, we will try to determine this angular dependence from video-tracking data. The distance dependence $g(d_{\alpha\beta}(t))$ has been specified in different ways. We will start with the **circular specification** of the interaction force,

$$\vec{g}(d_{\alpha\beta}) = A_\alpha e^{(R_\alpha + R_\beta - d_{\alpha\beta})/B_\alpha} \hat{d}_{\alpha\beta}, \quad (5.5)$$

where R_α and R_β denote the radii of pedestrians α and β . A_α reflects the strength of interaction, and B_α corresponds to the interaction range.

5.2.1 Improved Specifications of the Social-Force Model

Interactions among pedestrians are actually more complicated than suggested above. For example, it is known that the angle $\delta_{\alpha\beta}$ matters, at which two pedestrians α and β approach each other (which is given by $\cos(\delta_{\alpha\beta}) = \vec{e}_\alpha \cdot \vec{e}_\beta$). Apart from this, the step sizes and, therefore, the speeds matter as well. In the following, we will shortly discuss two anisotropic models of pedestrian interactions.

Elliptical specification I: In Ref. [57], a generalization of Eq. (5.5) was formulated, which assumed that the repulsive potential

$$V_{\alpha\beta}(b_{\alpha\beta}) = AB e^{-b_{\alpha\beta}/B} \quad (5.6)$$

is an exponentially decreasing function of $b_{\alpha\beta}$ with equipotential lines having the form of an ellipse directed into the direction of motion. The semi-minor axis $b_{\alpha\beta}$ is determined

by

$$2b_{\alpha\beta} = \sqrt{(\|\vec{d}_{\alpha\beta}\| + \|\vec{d}_{\alpha\beta} - v_\beta \Delta t \vec{e}_\beta\|)^2 - (v_\beta \Delta t)^2} \quad (5.7)$$

in order to take into account the length $v_\beta \Delta t$ of the stride (step size) of pedestrian β , where $v_\beta = \|\vec{v}_\beta\|$. The reason for this specification was that pedestrians require space for movement, which is taken into account by other pedestrians.

The repulsive force is related to the repulsive potential via

$$\vec{g}_{\alpha\beta}(\vec{d}_{\alpha\beta}) = -\vec{\nabla}_{\vec{d}_{\alpha\beta}} V_{\alpha\beta}(b_{\alpha\beta}) = -\frac{dV_{\alpha\beta}(b_{\alpha\beta})}{db_{\alpha\beta}} \vec{\nabla}_{\vec{d}_{\alpha\beta}} b_{\alpha\beta}(\vec{d}_{\alpha\beta}). \quad (5.8)$$

Considering the chain rule, $\|\vec{z}\| = \sqrt{\vec{z}^2}$, and $\vec{\nabla}_{\vec{z}}\|\vec{z}\| = \vec{z}/\sqrt{\vec{z}^2} = \hat{\vec{z}}$, this leads to the explicit formula

$$\vec{g}_{\alpha\beta}(\vec{d}_{\alpha\beta}) = A e^{-b_{\alpha\beta}/B} \cdot \frac{\|\vec{d}_{\alpha\beta}\| + \|\vec{d}_{\alpha\beta} - \vec{y}_{\alpha\beta}\|}{2b_{\alpha\beta}} \cdot \frac{1}{2} \left(\frac{\vec{d}_{\alpha\beta}}{\|\vec{d}_{\alpha\beta}\|} + \frac{\vec{d}_{\alpha\beta} - \vec{y}_{\alpha\beta}}{\|\vec{d}_{\alpha\beta} - \vec{y}_{\alpha\beta}\|} \right) \quad (5.9)$$

with $\vec{y}_{\alpha\beta} = v_\beta \Delta t \vec{e}_\beta$. For $\Delta t = 0$, we regain the expression of Eq. (5.5).

Elliptical specification II: Recently, a variant of this approach has been proposed [120, 80], assuming

$$2b_{\alpha\beta} = \sqrt{(\|\vec{d}_{\alpha\beta}\| + \|\vec{d}_{\alpha\beta} - (\vec{v}_\beta - \vec{v}_\alpha)\Delta t\|)^2 - \|(\vec{v}_\beta - \vec{v}_\alpha)\Delta t\|^2}. \quad (5.10)$$

The explicit force specification is obtained by using $\vec{y}_{\alpha\beta} = (\vec{v}_\beta - \vec{v}_\alpha)\Delta t$ in Eq. (5.9). The special feature of this approach is its symmetrical treatment of both pedestrians α and β . We would, however, like to mention that further velocity-dependent specifications of pedestrian interaction forces have been proposed in the past, for example, by Molnár [99, 49], Okazaki [105], Hoogendoorn [64], and Yu *et al.* [142].

5.3 Empirical Data

We have performed video tracking with the method described in Sec. 2.2, for the three videos depicted in Fig. 5.2. Two of the videos were recorded in natural environments in Budapest, Hungary, and the third one is from an experiment [51] conducted in Dresden, Germany.

As a result we obtain a number of datasets containing pedestrian trajectories. Let us now analyze these datasets and extract information of pedestrian behavior and interaction.

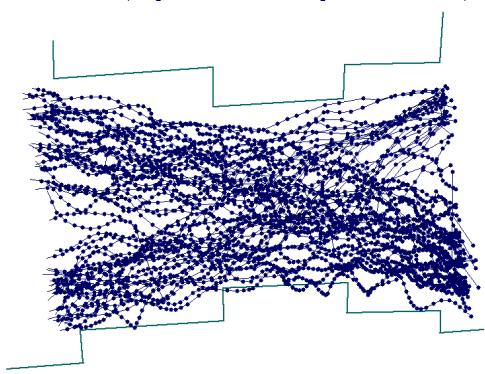
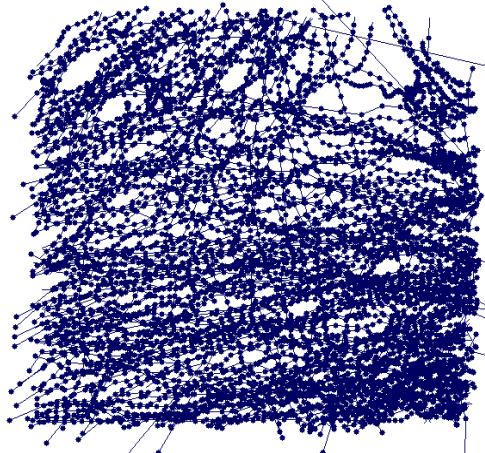
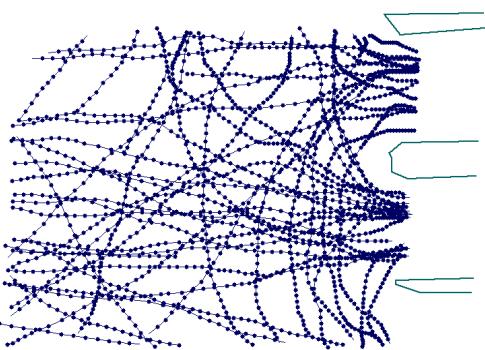


Figure 5.2: Snapshots of the video recordings, and corresponding trajectories, for the three video recordings used for parameter calibration. Top: (Video A) Entrance and exit area of two escalators, in Budapest, Hungary. Middle: (Video B) Free movement in a shopping mall in Budapest, Hungary. Bottom: (Video C) Pedestrian crossing experiment with students (see Ref. [51] for details), from Dresden, Germany. In the first two cases, pedestrians did not know they were recorded. Therefore, it can be assumed that their behavior was not influenced by the recordings.

5.3.1 Trajectory Data

The trajectories obtained from the video analysis consist of time series of observed positions $\vec{r}_\alpha^{obs.}(t_k)$, observed velocities $\vec{v}_\alpha^{obs.}(t_k)$, and observed accelerations $\vec{a}_\alpha^{obs.}(t_k)$ for all pedestrians α , and all discrete time samples, $t_k = k\Delta t$, $k = 0, 1, 2, \dots$.

Since videos are composed by time series of frames (images), the resulting trajectories datasets get a sampling time Δt ¹ corresponding to the frame rate of the videos. The videos are recorded in 12.5² FPS (frames per second), which gives a sampling time $\Delta t = 1/12.5$ s.

To avoid the boundary effects when a pedestrian suddenly enters or leaves the video-recorded area (and thus disappears from the data), we use linear extrapolation to define values for all pedestrians α in the dataset, before the time t_α^0 when the pedestrian enters the scene as well as after time t_α^1 when the pedestrian leaves the scene. The resulting (observed and extrapolated) trajectory for pedestrian α is defined as:

$$\vec{r}_\alpha(t) = \begin{cases} \vec{r}_\alpha^{obs.}(t_\alpha^0) + (t - t_\alpha^0)\vec{v}_\alpha^{obs.}(t_\alpha^0), & \text{if } t \leq t_\alpha^0 \\ \vec{r}_\alpha^{obs.}(t), & \text{if } t_\alpha^0 < t < t_\alpha^1 \\ \vec{r}_\alpha^{obs.}(t_\alpha^1) + (t - t_\alpha^1)\vec{v}_\alpha^{obs.}(t_\alpha^1), & \text{if } t \geq t_\alpha^1 \end{cases} \quad (5.11)$$

5.4 Hybrid Data-Analysis Approach

Fundamental quantities such as walking speeds, densities and time headways can be directly extracted from the trajectory datasets, by means of regressions and histograms.

More complex quantities however, such as interaction behaviors, can not directly be extracted from the data. Therefore a hybrid method [80] is used, fusing empirical trajectory data and microscopic simulation of pedestrian movement (see Fig. 5.3).

To each observed pedestrian in the dataset, a corresponding virtual pedestrian is assigned in the simulation domain. A simulation is initiated and runs for $T = 1.5$ seconds, in which one pedestrian α is moved according to a simulation with the social-force model [57, 53, 51] (see Sec. 5.2), while the others are moved according to their trajectories as extracted from the videos.

The reason why $T = 1.5$ seconds is used, is elaborated and explained in detail in Sec. 5.5.

¹Note that this Δt is different from the social-force-model parameter with the same name.

²Actually, the videos were originally recorded with 25 FPS, but every second frame was dropped before the analysis.

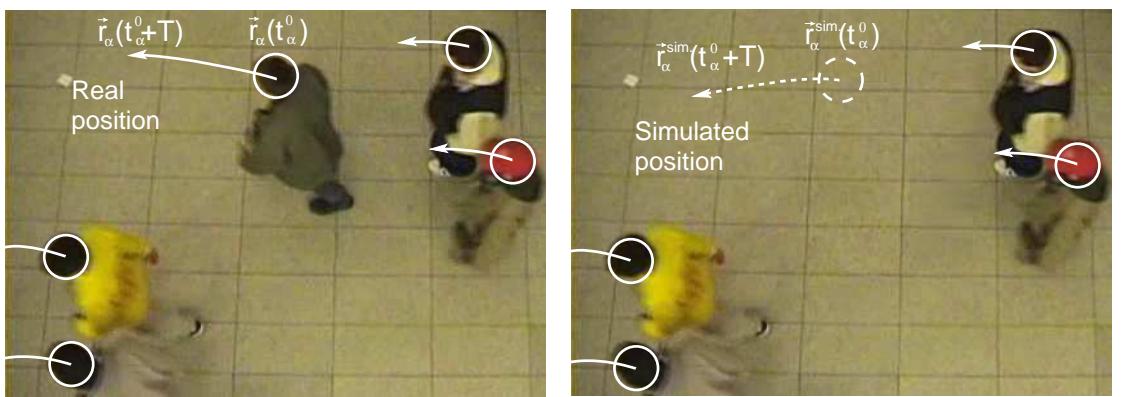


Figure 5.3: Illustration of the hybrid approach, used for fusing empirical trajectory data with microscopic simulations. Left: All pedestrians (solid circles) follow the empirical trajectories. Right: For each pedestrian α (dashed circle) and for different starting times t_α^0 , a simulation is run during T seconds. One can then obtain a fitness value of the simulation by comparing the actual position to the simulated position.

The forecasting procedure is repeated for every pedestrian α and for several different starting times t_α^0 , using a fixed parameter set for the social-force model.

Each simulation run is performed according to the following scheme:

1. Assign a desired speed v_α^0 to each pedestrian. In our simulations, we specify the desired speed by first determining all individual speeds $\|\vec{v}_\alpha(t)\|$ for $t = t_\alpha^0 + k\Delta t$ along the trajectory of α . Then we sort these speeds, and finally choose the speed corresponding to a certain percentile. Note that we assume that a pedestrian's desired speed is constant.
2. Define a starting point and calculate the state (position \vec{r}_α , velocity \vec{v}_α , and acceleration $\vec{a}_\alpha = d\vec{v}_\alpha/dt$) of each pedestrian α .
3. Assign a desired target location for each pedestrian. We assume that it corresponds to the point at the end of the trajectory.
4. Given the tracked motion of the surrounding pedestrians β ($\neq \alpha$), simulate the trajectory of pedestrian α based on the social-force model over a time period T , starting at the actual location $\vec{r}_\alpha(t_\alpha^0)$.

After each run, we determine the squared distance error

$$E = \left\langle \frac{1}{T} \int_{t_\alpha^0}^{t_\alpha^0 + T} \|\vec{r}^{sim.}(t) - \vec{r}(t)\|^2 dt \right\rangle_{\alpha, t_\alpha^0} \approx \left\langle \frac{1}{N} \sum_{i=0}^{N-1} \|\vec{r}^{sim.}(t_\alpha^0 + i\Delta t) - \vec{r}(t_\alpha^0 + i\Delta t)\|^2 \right\rangle_{\alpha, t_\alpha^0} \quad (5.12)$$

which is averaged over all pedestrians α and different starting times t_α^0 . The summation is made for $N = (t_\alpha^1 - t_\alpha^0)/\Delta t + 1$. The error function differs in two important ways from the one used in Ref. [80]. First, we integrate the error over the whole time evolution T rather than only measuring the error once after the time has progressed T seconds. Second, the error is squared, which favors small fluctuations from the observed trajectory, but strongly punishes eventual large deviations. The square is introduced for the optimization to grasp the general shapes of the trajectories and to prevent over-fitting.

After averaging the squared distance errors over the pedestrians α and starting times t_α^0 , the value is taken as the “*fitness*” of the parameter set used in the pedestrian simulation.

Hence, the best possible value of the fitness is 0, and increasing deviations from the real pedestrian trajectories give increasing fitness values. Therefore the “*fitness*” is actually the *error*. For examples of comparisons of real trajectories with simulated trajectories and their corresponding fitness values, see Fig. 5.4.

With the scheme detailed above, we can now implement a feedback on what “*fitness*” a certain pedestrian model together with a certain parameter set has. With this feedback we can run a genetic algorithm [96] which finds the parameter set with the best “*fitness*”. The genetic algorithm³ (see *Algorithm 2 on page 79*) is applied to a “*population*” of 100 individuals, each characterized by the parameters used in the current model.

In step 4, a fraction $F_1 = (1 + \tanh(rnd[-2.5, 2.5]))/2$ of the parameter values are taken from the 1st parent, and a fraction $1 - F_1$ of the parameter values are taken from the 2nd parent, where $rnd[-2.5, 2.5]$ gives uniformly distributed random values on the interval $[-2.5, 2.5]$. The non-linearity, provided by the tanh function is important for the following reason: For multi-modal fitness surfaces, the equal (50%/50%) combination of two good solutions is very likely to lead to a solution that is worse than any of the two original solutions. Therefore, by specifying the factor F_1 in a non-linear way, we take primarily one solution which is then “influenced” by another solution. After the

³The reason for the name *genetic algorithm* is that a “*population*” is used, containing different parameter sets, reminding of a “*gene pool*”, where bad genomes are replaced by better ones, from re-combinations and mutations. Thereby, genomes with good fitness survive and are used for “*reproduction*” and bad genomes are removed, which drives the population towards better fitness.

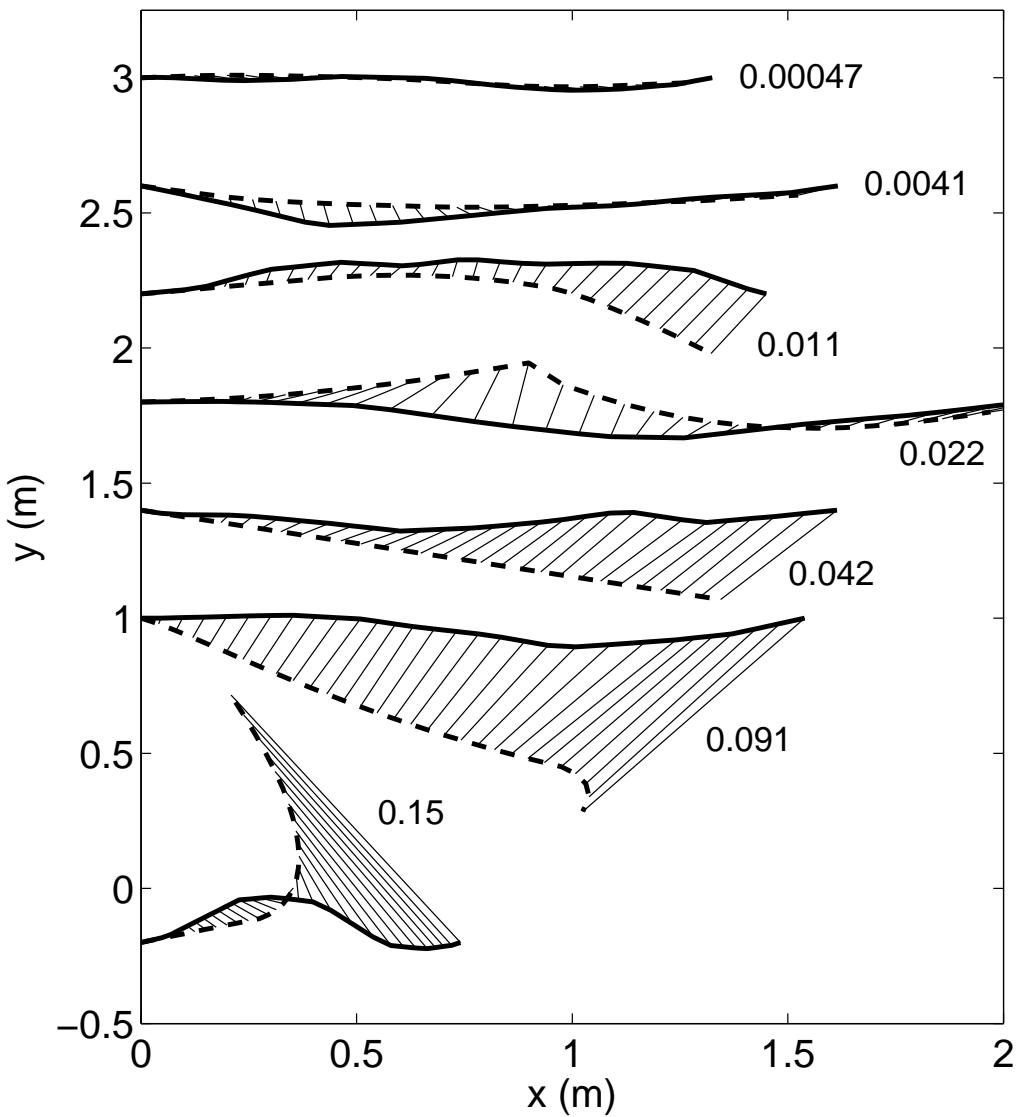


Figure 5.4: Examples of empirical trajectories (thick solid lines) compared to their corresponding simulated trajectories (dashed lines) and their resulting fitness values measured by Eq. (5.12). The distance errors $\vec{r}^{sim.}(t) - \vec{r}(t)$ are displayed as thin solid lines. See Tab. 5.2 for the corresponding fitness values after model calibration.

Algorithm 2 Genetic Algorithm

1. Initiate the population by assigning random parameter values within pre-defined intervals to all individuals.
 2. Choose four individuals i, j, k, l , at random.
 3. Run the simulation for each of the four individuals, and compute their “fitness” values f_i, f_j, f_k, f_l , with the given objective function (Eq. 5.12).
 4. Remove the two individuals with the highest error E (Eq. 5.12) from the population and replace them by two new individuals that are created by combination and random mutation of parameter values of the two individuals with the best fitness (lowest error) values.
 5. Goto (2)
-

combination has been performed, two different kinds of mutations are applied:

- Entirely new values are randomly assigned to 10% of the parameters.
- Small adjustments are made on 30% of the parameters, by increasing or decreasing the parameter value by a uniformly distributed random value in the interval [-20%, 20%].

5.4.1 Distance-Dependence of Pedestrian Interactions

The genetic algorithm can be used to determine interaction laws without using pre-specified interaction functions. To show this, we start by using an isotropic angular-dependence $w(\varphi_{\alpha\beta}(t)) = 1$, and then finding the distance-dependent function $\vec{g}(d_{\alpha\beta}(t))$ for pedestrian interaction, which is the second part of Eq. (5.3). First, we assume that the direction of \vec{g} corresponds to $\vec{r}_\alpha(t) - \vec{r}_\beta(t)$. Second, since we do not want to assume a certain functional form, the magnitude of \vec{g} is adjusted at given distances $d_k = kd_1$ (with $k \in \{0, 1, 2, 3, \dots\}$ and $d_1 = 0.5$ m) with our genetic algorithm. The optimization is applied on the three different datasets shown in Fig. 5.2, and the resulting curve is presented in Fig. 5.5 by its average value as well as error bars corresponding to the standard deviations from the three different curves, obtained from each of the datasets. According to our results, the empirical dependence of the force with distance can be well fitted by an exponential decay, as hypothesized before [57].

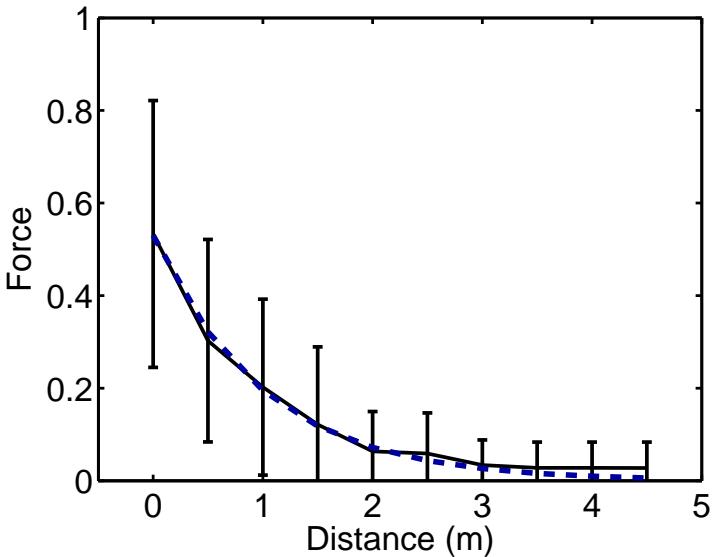


Figure 5.5: Empirically determined distance dependence of the interaction force between pedestrians. An exponential decay fits the empirical data well. The dashed fit curve is made with $F = A \exp[-(d_{\alpha\beta} - R_\alpha - R_\beta)/B]$ where $A = 0.53$ and $B = 1.0$. $d_{\alpha\beta} - R_\alpha - R_\beta$ is the net distance between pedestrians α and β . The solid curve corresponds to the average values of three curves obtained by our genetic algorithm when applied to the three datasets illustrated in Fig. 5.2. The error bars correspond to the standard deviation of the three curves. Note that, even though an exponential decay fits the data well, several other functional forms would fit within the error bars as well.

5.4.2 Angular-Dependence of Pedestrian Interactions

Now, we use the force specification of $\vec{g}(d_{\alpha\beta})$ found by our optimization method and look at the angular dependence of pedestrian interaction (first part of Eq. (5.3), i.e. $w(\varphi_{\alpha\beta}(t))$). Since we want to find the best function, without having to pre-specify its shape, we choose to model the angular-dependence function by a polygon with n edges in the directions $\varphi_i = 2\pi i/n, i \in [1, n]$. The distance of the curve in direction φ_i from the origin represents the (relative) angular interaction strength $w(\varphi_i)$ with a maximum value of 1. To get a smooth curve, linear interpolation is applied between each of the edges.

Next, we run the genetic algorithm to find the interaction strengths $w(\varphi_i)$ giving the highest fitness value. By doing so, Fig. 5.6 shows the angular dependence of the interaction strength determined from pedestrian trajectories from different datasets. Apparently, pedestrians are only sensitive to what happens inside an 180-degrees angle in front of them, which roughly corresponds to the visually perceived area (considering also head movements).

5.4.3 Number of Interactions

Until now, we have followed the lines of thought of the original social-force model [57, 53], where forces from neighboring pedestrians (and boundaries) are superposed to create the resultant force. However, the validity of this approach has not been tested, and we will find evidence suggesting a modification of it:

1. Assume that a pedestrian α is standing at the last position of a queue. It is reasonable to assume that he/she will adapt the distance only to the closest pedestrian in the queue. When interaction with all pedestrians are added up however, he/she will feel an increasingly strong force, the longer the queue is. The strength of this effect depends on how fast the forces decay as a function of distance.
2. The superposition strategy assumes that a pedestrian is able to keep track of and “feel” the presence and relative orientation of all the pedestrians in the whole scenario. It would be reasonable to assume that a pedestrian rather keeps track of a certain maximum number N of neighboring pedestrians only (the closest ones), and temporarily neglects the rest.

In order to test our hypothesis that a pedestrian does not interact with all pedestrians around him/her, but only with a maximum number N , we study our data once more.

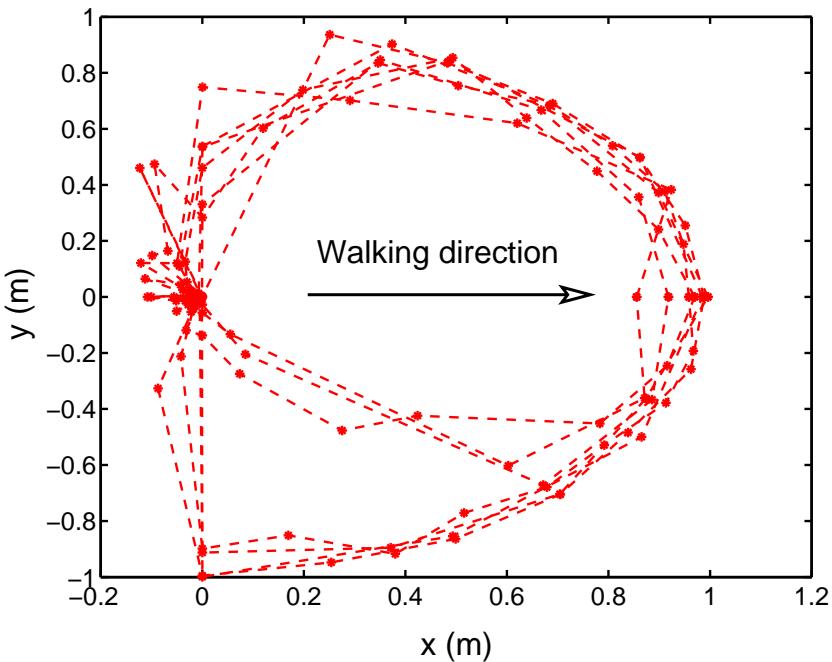


Figure 5.6: Angular dependence of the influence of other pedestrians. The direction along the positive x axis corresponds to the walking direction of the pedestrian, y to the perpendicular direction. Each curve corresponds to the result from one optimization run.

Unfortunately, our three videos (shown in Fig. 5.2) do not cover large enough areas to make them feasible for finding N . For this reason, we are using the trajectory data analyzed in Ref. [69]. These data correspond to a walking experiment, where two streams of pedestrians intersect at a 90-degrees angle.

Five separate parameter optimization runs are made for $N = 1, 2, 3, \dots, 10$. The mean value \pm one standard deviation of these five runs are shown in Fig. 5.7. Naturally, the fitness value (i.e. the error) is initially decreasing with increasing N , until $N = 7$.

However, the data suggest that the fitness may even *increase* for $N > 8$, indicating that the pedestrians in the video typically react to their 7 to 8 closest neighbors only.

When taking into consideration that the social forces are significantly anisotropic, which makes pedestrians react mainly to what happens in front of them, we can reformulate our result to: *Pedestrians mainly react to their 3–4 closest pedestrians in front of them.*

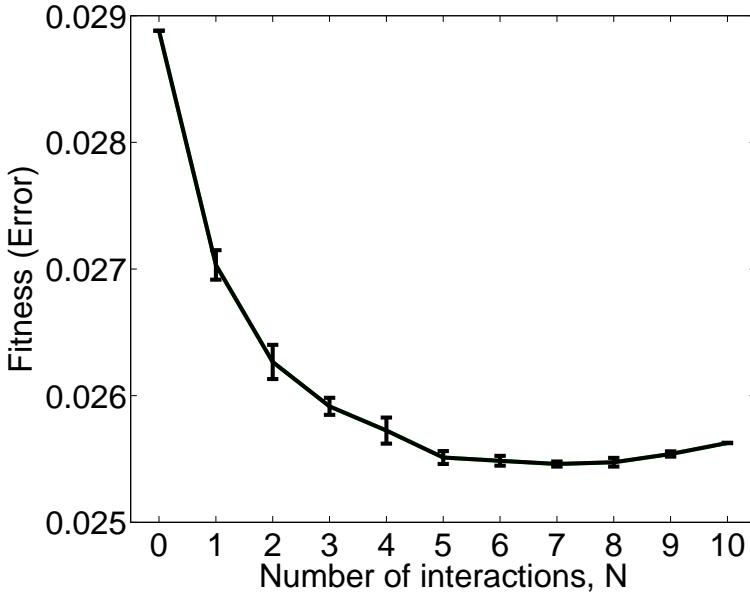


Figure 5.7: Result of running parameter optimization for fixed values of $N = 1, 2, \dots, 10$. Naturally, the fitness value (i.e. the error) is initially going down by an increasing N . However, when reacting to more than 7–8 people, the fitness is slightly *increasing* again. This indicates that the pedestrians in the video typically react to their 7 to 8 closest neighbors only. The curve corresponds to the average value from five separate optimization runs, and the error bars to one standard deviation.

5.4.4 Delays

As has been pointed out before [124, 66], it may be important to introduce an explicit delay in pedestrian models, to take into account the delay of the human brain, when it comes to (conscious) decision-making.

In order to investigate this, we again use our genetic algorithm with the added parameter T^d , representing the delay in seconds of the model. Then Eq. (5.3) is substituted by:

$$\vec{f}_{\alpha\beta}(t) = w(\varphi_{\alpha\beta}(t - T_\alpha^d)) \vec{g}(d_{\alpha\beta}(t - T_\alpha^d)). \quad (5.13)$$

The individual reaction time T_α^d is found by optimizing the fitness of each pedestrian α , for the model (Eq. 5.13). The result is shown in Fig. 5.8.

Interestingly, the probability-density function of the delay curve is bi-modal. The first peak occurs at lower times than the typical response times to visual or acoustic cues [139, 137]. Therefore, this peak must relate to *anticipated* movements of the surrounding pedestrians. The second peak at around 0.45 s occurs at times which are significantly larger than the previously mentioned response times, but also lower than response times involving conscious reactions [131, 40]. Therefore, we conclude that this second peak corresponds to a subconscious response, which is more complex than an automatic response to anticipated movements. In fact, it has been shown that reactions with significantly longer reaction times occur for situations with more than one possible response (choice reaction time) as well as for reactions to more complex cues (recognition reaction time) [30].

We interpret the bi-modality as follows: When the surrounding pedestrians act in a way that is easy to predict, extrapolation allows to anticipate their behaviors, while a delayed reaction results in cases of unexpected behaviors.

Now, since pedestrians have a finite delay in their avoidance manoeuvres, it would be natural to assume that they have come up with a strategy to compensate for this delay. If this is the case, rather than pedestrians reacting to the *current* locations of the other pedestrians at time t , pedestrians would react to the extrapolations of the other pedestrians' locations, at time $t + T^d$. Therefore we have to substitute $\vec{r}_\beta(t)$ by $\vec{r}_\beta(t) + \vec{v}_\beta(t)T^a$ in the social-force model. This substitution corresponds to a linear extrapolation T^a forward in time, based on the current position and velocity of pedestrian β at time t . T^a is the anticipation time, with a value that is supposed to be similar to the one of T^d . However, in the *Elliptical II* specification of the social-force model (see Sec. 5.2.1), we have already introduced anticipation by the velocity-dependence of the

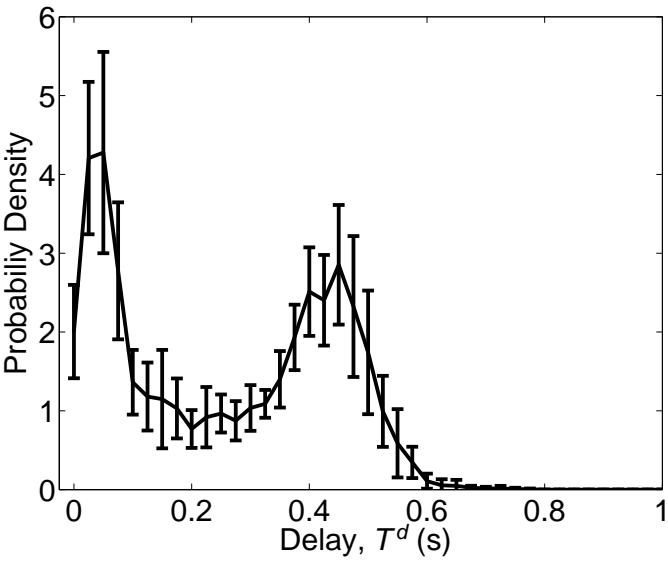


Figure 5.8: Probability-density function of the delay parameter T^d . The error bars correspond to one standard deviation. Interestingly, the probability-density function is bi-modal. When the neighboring pedestrians are acting in a way that is easy to predict, extrapolation allows to anticipate their behaviors, while a delayed reaction results in cases of unexpected behaviors.

forces. Therefore we will not introduce an additional anticipation in the model.

5.5 Calibrating the Social-Force Model

Parameter	Meaning	Unit
A	Interaction strength	m/s^2
B	Distance scaling	m
N	Number of interactions	
τ	Relaxation time	s
ΔT	Anticipation time	s
SP	Speed percentile	%
λ	Anisotropy constant	
T^d	Delay	s

Table 5.1: A summary of the parameters used by the social-force model.

Let us now turn to the question of how the parameters of the social-force model (listed in Table 5.1) can be calibrated with our hybrid method detailed in Sec. 5.4. Calibration of the model parameters with empirical data have been made in the past with vehicle data [33, 131] and recently also for pedestrians [80, 67].

Similar to what has been done in Ref. [80], we use the hybrid method detailed in Sec. 5.4, in order to find the parameter values giving the best “fitness” value, as defined by Eq. (5.12).

The remaining lowest fitness value reflects the stochasticity and heterogeneity of individual behaviors.

As a forecasting time, $T = 1.5$ seconds is used, which turns out to be a good value because for low forecasting times ($T < 1.0$ s) any combination of low A or B values give fairly good fitness values. This is interpreted as follows: *For a very short forecasting time, a zero-acceleration model gives realistic results, where as for larger forecasting times ($T > 1.0$ s), a more sophisticated model is needed in order to fit the trajectories.* See Fig. 5.9 for the fitness surface, spanned by the parameters A and B , for four different values of T . Even larger values of T would exclude too many of the empirical trajectories, since many of the pedestrians are not present in the video for longer time than a few seconds.

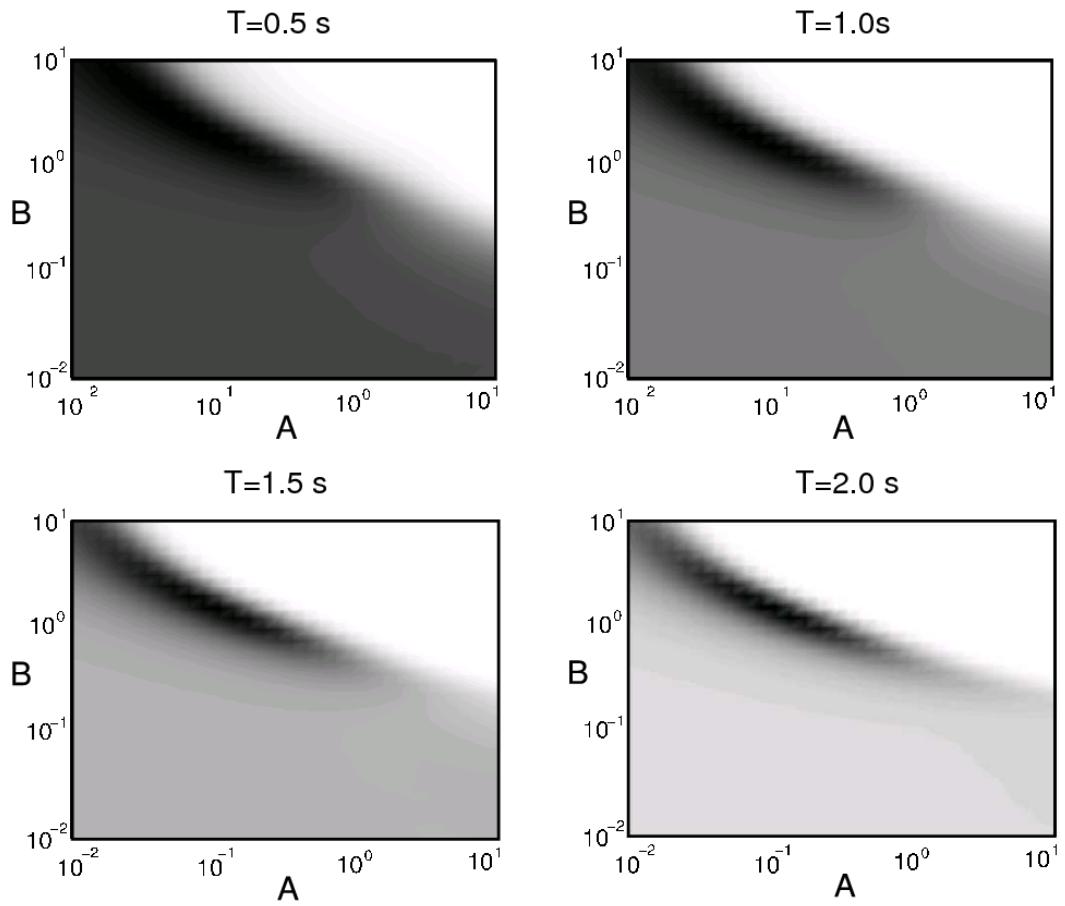


Figure 5.9: Fitness surface, spanned by the parameters A and B , for different values of the forecasting time T . The darker shades correspond to the best fitness values. For low values of the forecasting time T , any combination of low A or B values give fairly good fitness values. For this reason, we choose the value $T = 1.5$ s for our further analysis. Even larger values of T would exclude too many of the empirical trajectories, since many of the pedestrians are not present in the video for longer time than a few seconds.

5.5.1 Parameter Calibration

Our first approach to parameter calibration is to assume a *homogeneous* population, i.e. assume that all pedestrians share the same parameter set. By doing so, and running the parameter optimization for datasets A, B, and C, we get the resulting fitness values displayed in Table 5.2 and parameter values displayed in Table 5.3.

Apparently there is a very small variation of fitness values between the different model specifications⁴. For this reason, we make a second parameter optimization by assigning a unique set of parameters to each pedestrian (referred to as *heterogeneous parameters*). As the second part of Table 5.2 shows, the separation is larger between the different model specifications when heterogeneous parameters are used.

5.5.2 Parameter Analysis

As in most models, the parameters of the social-force model are not orthogonal to each other, meaning that all parameters are correlated with each other to different degrees.

By assuming isotropic forces $w(\varphi_{\alpha\beta}) \equiv 1$ and by varying parameters A and B in the intervals $[0, 10]$ and $[0, 10]$ we get a resulting fitness-surface spanned by parameters A and B , as shown in Fig. 5.10.

Apparently there is not one outstanding $\{A, B\}$ pair with the best fitness. Rather the best values are distributed along a curve.

In an attempt to investigate why the values along this curve are good ones and what (if anything) they have in common, we pick a few $\{A, B\}$ pairs along the curve and plot the force specified in Eq. (5.5). As can be seen in Fig. 5.10, for the two videos with a bit higher density (videos B and C) the distance-dependent forces all agree at one distance. We interpret this result in the following way: *There is one main decision point at a certain distance d where pedestrians decide to change their velocity in order to evade the other pedestrian.* However, this decision point is not seen in video A, where the density is low, and it is also at a shorter distance in video C than in video B, which can be explained by the significantly lower walking speeds in video C compared to video B. Therefore, we can interpret the decision point to be located at a certain time headway rather than at a certain distance.

⁴The reason for the larger separation of fitness values in Ref [80] is partly because less parameters were optimized in that study (giving a stronger dependence of the model specification rather than on parameter values). Second, in Ref. [80] the *median* rather than the *mean* fitness values was used, which makes the method less sensitive to outliers.

Model	N	Angular dependence	Fitness mean	Fitness dataset A	Fitness dataset B	Fitness dataset C
Homogeneous parameters						
Circular	Variable	Variable λ	0.0390	0.0335	0.0376	0.0460
Elliptical I	Variable	Variable λ	0.0387	0.0341	0.0365	0.0455
Elliptical II	Variable	Variable λ	0.0383	0.0328	0.0365	0.0456
Heterogeneous parameters						
Circular	Variable	Variable λ	0.0142	0.0064	0.0128	0.0234
Elliptical I	Variable	Variable λ	0.0114	0.0041	0.0099	0.0201
Elliptical II	Variable	Variable λ	0.0106	0.0039	0.0086	0.0194
Circular	Variable	$\lambda = 1$	0.0151	0.0066	0.0137	0.0250
Elliptical I	Variable	$\lambda = 1$	0.0125	0.0052	0.0106	0.0219
Elliptical II	Variable	$\lambda = 1$	0.0119	0.0046	0.0096	0.0214
Circular	Variable	Half circle	0.0151	0.0069	0.0132	0.0251
Elliptical I	Variable	Half circle	0.0121	0.0045	0.0105	0.0212
Elliptical II	Variable	Half circle	0.0118	0.0047	0.0101	0.0206
Circular	∞	Variable λ	0.0157	0.0071	0.0142	0.0257
Elliptical I	∞	Variable λ	0.0119	0.0044	0.0103	0.0209
Elliptical II	∞	Variable λ	0.0115	0.0044	0.0093	0.0208
Circular	∞	$\lambda = 1$	0.0166	0.0075	0.0152	0.0272
Elliptical I	∞	$\lambda = 1$	0.0131	0.0050	0.0116	0.0227
Elliptical II	∞	$\lambda = 1$	0.0123	0.0056	0.0101	0.0211
Circular	∞	Half circle	0.0161	0.0073	0.0146	0.0266
Elliptical I	∞	Half circle	0.0130	0.0046	0.0113	0.0231
Elliptical II	∞	Half circle	0.0120	0.0047	0.0103	0.0211

Table 5.2: Resulting fitness values for different parameter-optimization runs, by assuming *homogeneous* parameters as compared to *heterogeneous* parameters, i.e. a unique parameter set per pedestrian. Each row corresponds to a unique combination of the three first columns (*model*, *N* and *angular dependence*). The resulting fitness values are shown as a mean fitness from the three datasets (videos) A, B and C shown in Fig. 5.2 as well as the average fitness from each of the three datasets.

Model	N	Angular dependence	A_s	B_s	τ	Speed percentile		ΔT
						53	42	
Circular	10	$\lambda = 0.00$	0.23	0.70	0.61	53	-	
Elliptical I	12	$\lambda = 0.00$	0.53	0.48	0.60	53	1.27	
Elliptical II	12	$\lambda = 0.00$	0.33	0.55	0.60	52	0.97	
Circular	1	$\lambda = 1$	0.21	0.88	0.57	42	-	
Elliptical I	1	$\lambda = 1$	0.44	0.37	0.57	42	0.95	
Elliptical II	19	$\lambda = 1$	0.53	0.29	0.55	42	1.52	
Circular	10	Half circle	0.21	0.82	0.61	53	-	
Elliptical I	10	Half circle	0.23	0.85	0.60	54	2.57	
Elliptical II	11	Half circle	0.21	0.89	0.60	53	1.07	
Circular	∞	$\lambda = 0.00$	0.32	0.51	0.60	52	-	
Elliptical I	∞	$\lambda = 0.00$	0.55	0.46	0.60	53	1.26	
Elliptical II	∞	$\lambda = 0.00$	0.25	0.59	0.60	52	1.27	
Circular	∞	$\lambda = 1$	0.26	0.37	0.57	42	-	
Elliptical I	∞	$\lambda = 1$	0.35	0.32	0.57	43	0.87	
Elliptical II	∞	$\lambda = 1$	0.45	0.32	0.57	42	1.61	
Circular	∞	Half circle	0.25	0.63	0.62	53	-	
Elliptical I	∞	Half circle	0.22	0.70	0.61	54	2.59	
Elliptical II	∞	Half circle	0.21	0.72	0.62	53	1.06	

Table 5.3: Resulting parameter values for different parameter-optimization runs. Each row corresponds to a unique combination of the three first columns (*model*, *N* and *angular dependence*).

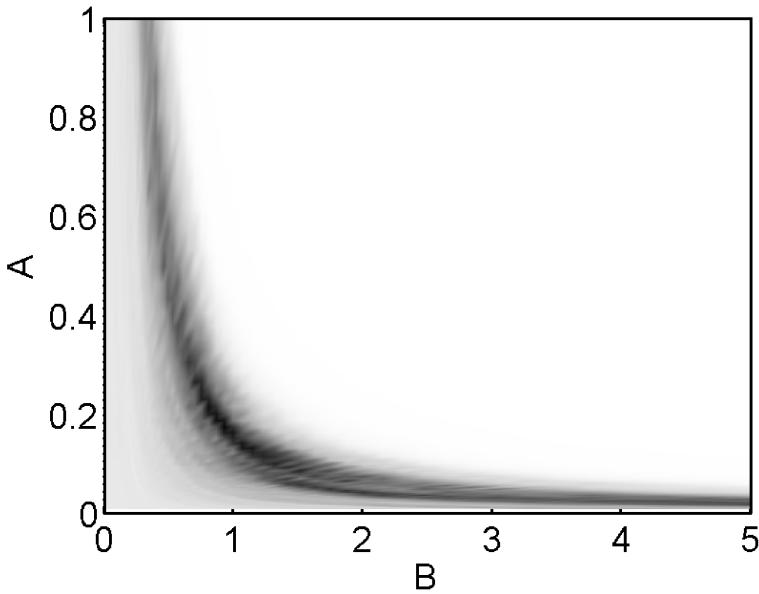


Figure 5.10: Fitness values for different parameter values A and B , when pedestrian trajectories are simulated with the social-force model and the force specification in Eq. (5.5). The darker the shade is, the better the fitness, for the corresponding combination of A and B values.

Next, let us look at the resulting parameter values for the best performing social-force model specification, which is the *Elliptical II* specification with a variable parameter N and a variable λ . The resulting parameter values of our three datasets are put after each other in matrix P , where each row corresponds to a pedestrian and each column corresponds to a parameter. The covariance $C = \text{cov}(P)$ is computed and in order to analyze how different parameters correlate with each other, we compute the correlation coefficient $R(i, j)$ which measures how well parameter i correlates with parameter j . It is defined via the covariance C as

$$R(i, j) = \frac{C(i, j)}{\sqrt{C(i, i)C(j, j)}}. \quad (5.14)$$

The correlation coefficient spans from perfect correlation $R(i, j) = 1$ (rows i and j differ only by a scaling constant) via zero correlation $R(i, j) = 0$ ($C(i, j) = 0$) to perfect de-correlation $R(i, j) = -1$, which means that columns i and j are orthogonal.

The resulting correlation coefficients $R(i, j)$ are shown in Tab. 5.4 with exceptional values highlighted in boldface. We see for example that parameter A correlates strongly

negatively with parameter B , which has already been investigated in quite some detail (see Fig. 5.10). The negative correlation between parameters A and T^d can be understood by different walking personalities, i.e. and active/aggressive walker style (low T^d and high A) and a more relaxed walking style (high T^d and low A).

The negative correlation between B and ΔT can be interpreted as the possibility that pedestrians use two different heuristics for prevent crashing into other pedestrians:

- *Heuristic I:* Using a large distance-scaling B of the repulsive forces prevents coming too close to any of the surrounding pedestrians.
- *Heuristic II:* Using a large anticipation time ΔT will prevent crashing into pedestrians whose extrapolated position will interfere with one's own planned path.

The latter heuristic is more sophisticated and therefore requires a higher cognitive load in order to extrapolate the surrounding pedestrians' paths and look for potential path interferences ΔT seconds later. The first heuristic, on the other hand, requires a lower cognitive load, but is less predictive and exact since it reacts only to positions of surrounding pedestrians (and neglect their velocities).

Next, we see that there is a strong positive correlation between relaxation time τ and delay T^d . This is natural since relaxed pedestrians tend to have both a longer relaxation time as well as a larger delay, where more active pedestrians on the other hand tend to have both a smaller relaxation time as well as a lower delay.

Finally, we see that the anisotropic parameter λ correlates negatively both with the anticipation time ΔT as well as the walking speed. This means that pedestrians with a low λ (strong anisotropy) tend to walk faster and/or having a larger anticipation time ΔT .

The relative fitness decline is investigated by replacing personal parameters with global parameters. First, a unique parameter set per pedestrian is found, which results in the overall average fitness f^0 . Then, one of the parameters at a time is replaced by the value found when using one homogeneous parameter set shared by all pedestrians. By replacing parameter i for all pedestrians, the new overall mean fitness f_i is found. Finally, the relative fitness decline $(f_i - f^0)/f^0$ is measured. With this quantity, one can measure the individual variability of each parameter. A small relative fitness decline of parameter i means that either the fitness is not strongly dependent on this parameter, or that most pedestrians have similar values of this parameter. A large relative fitness decline of parameter i , on the other hand, means that there is a wide distribution of

A	B	N	τ	λ	Speed	ΔT	T^d	
-0.03	0.02	0.05	0.06	-0.05	-0.08	-0.07	-0.00	Fitness
-0.55	-0.13	-0.10	0.00		-0.14	0.03	-0.22	A
	-0.03	-0.08	0.07		0.04	-0.13	0.02	B
		0.17	-0.06		0.11	0.06	0.09	N
			0.15		0.03	-0.04	0.24	τ
				-0.20	-0.15	0.05		λ
					-0.10	-0.02		Speed
						0.03		ΔT

Table 5.4: Correlation coefficients R for the resulting parameter values from a parameter optimization of the *Elliptical II* specification of the social-force model for the three datasets A, B and C. Exceptional values are highlighted in boldface.

values of parameter i over all pedestrians and also that the fitness depends strongly on this parameter. What we can learn from this is which parameters are shared among all pedestrians, and which parameters have unique values for each pedestrian, reflecting something like *personality*. As can be seen in Tab. 5.5, the distance scaling parameter B and the interaction strength parameter A are the most important personality parameters.

Parameter	Relative fitness decline
Distance scaling B	161%
Interaction strength A	92%
Relaxation time τ	62%
Number of interactions N	38%
Anticipation time ΔT	34%
Speed	31%
Anisotropy λ	15%
Delay T^d	13%

Table 5.5: The relative fitness decline is investigated by replacing personal parameters with global parameters. The relative fitness decline gives the individual variability of each parameter. A small relative fitness decline of parameter i means that either the fitness is not strongly dependent on this parameter, or that most pedestrians have similar values of this parameter. A large relative fitness decline of parameter i , on the other hand, means that there is a wide distribution of values of parameter i over all pedestrians and also that the fitness depends strongly on this parameter. What we can learn from this is which parameters are shared among all pedestrians, and which parameters have unique values for each pedestrian, reflecting something like *personality*.

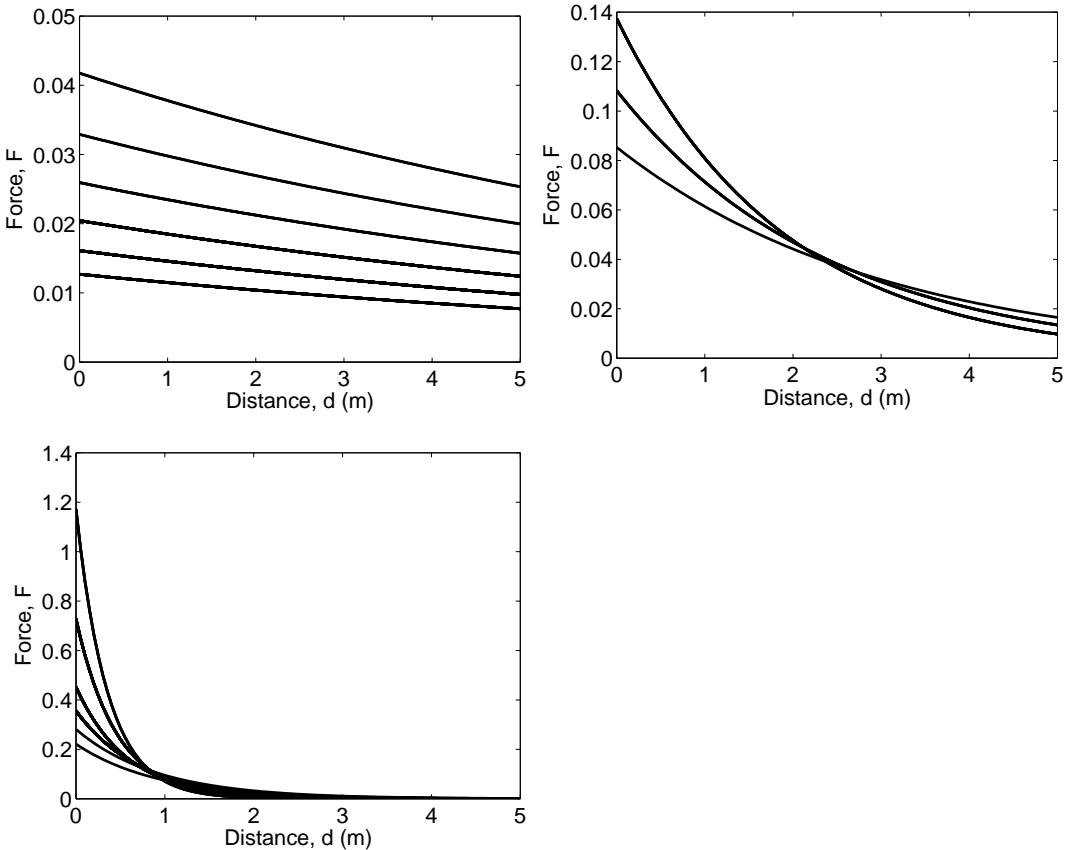


Figure 5.11: Top left: Dataset A. Top right: Dataset B. Bottom left: Dataset C. By plotting the curves with Eq. (5.5) and the $\{A, B\}$ -parameter pairs from Fig. 5.10 with the best fitness, we see that, for datasets B and C, the curves all agree in one point. We interpret this point as the main decision point, at where pedestrians decide to evade the others. Since the decision point is located at a larger distance in dataset B than in dataset C and the fact that the walking speeds are significantly higher in dataset B compared to dataset C, we can conclude that the decision point is located at a certain time headway rather than at a certain distance. The reason why the curves do not intersect in dataset A can be understood by the low density in this dataset, which makes the influence of evading manoeuvres negligible in comparison to other movements.

5.6 Simulation Results

In order to illustrate how the force field from the fellow pedestrian β is influencing pedestrian α , see Fig. 5.13. When two pedestrians are approaching each other at an 180-degrees angle, the elliptical specification results in smoother evading manoeuvres compared to the circular specification, which is illustrated in Fig. 5.12.

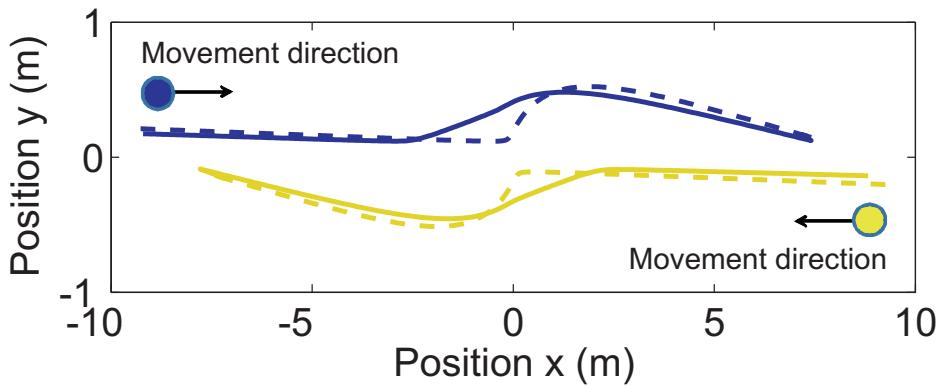


Figure 5.12: Resulting trajectories from a simulation of two pedestrians, who are approaching each other at an 180-degrees angle. The simulation is carried out with the social-force model, with two different model specifications. The *Elliptical II* specification (solid lines) results in smoother evading manoeuvres as compared to the *Circular* specification (dashed lines). The blue trajectory corresponds to a pedestrian walking into the right direction, and the yellow trajectory corresponds to a pedestrian walking into the left direction.

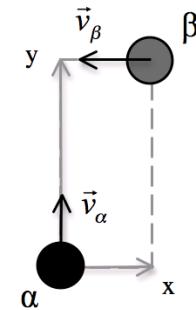
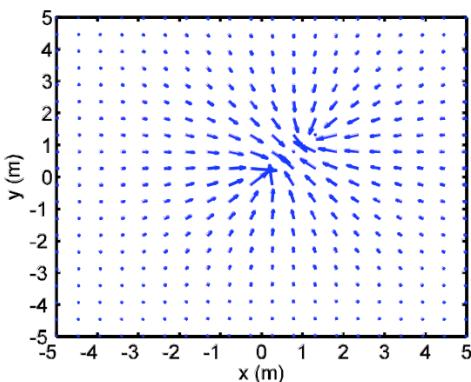
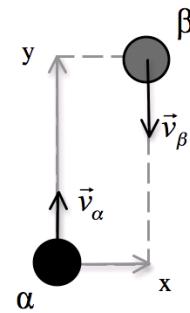
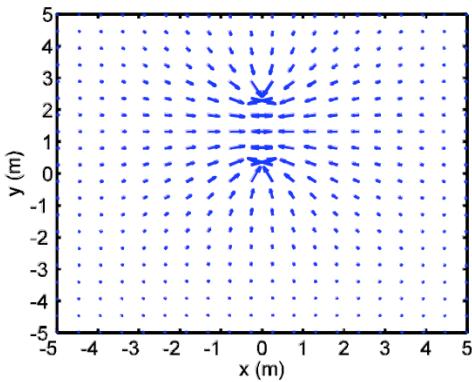
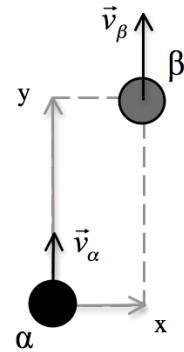
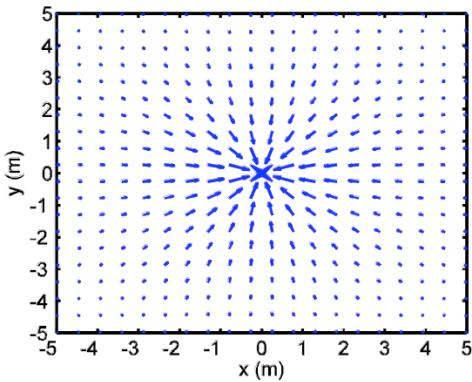


Figure 5.13: The force fields resulting from the *Elliptical II* specification of the social-force model. The force fields show how pedestrian α (located in the origin) is influenced by pedestrian β . Note that, in the top most figure, when $\vec{v}_\alpha = \vec{v}_\beta$, the *Elliptical II* specification is reduced to the *Circular* specification.

Given the social-force model from Sec. 5.2 that has been calibrated with empirical trajectory data with the method from Sec. 5.3, we now want to assess its ability to reproduce the various self-organization phenomena from Chapter 3.

We use the *Elliptical II* specification of the social-force model together with the parameters from Table 5.3. Since these phenomena all occur during high crowd density, we need to add an additional force which reflects physical interactions (as in Ref. [53]). This is done by adding an additional short-ranged force specified by Eq. (5.5) with $A = 3 \text{ m/s}^2$, $B = 0.2 \text{ m}$, $R = 0.15 \text{ m}$, and with an angular dependency specified in Eq. (5.4) with $\lambda = 0.12$.

We start by simulating a corridor with a bi-directional flow of pedestrians, where half of them walk into each of the two directions. The snapshot in Fig. 5.14 (top left) shows how the pedestrians self-organize into lanes.

Next, when a bottleneck is put in the middle of the corridor, waiting crowds form at each of the sides. Since the bottleneck is too narrow to allow simultaneous flows into both directions, pedestrians self-organize into oscillations of the flow direction, which can be seen in Fig. 5.14 (right).

When two pedestrian streams are intersecting at an angle different from 180 degrees, a clearly visible pattern called *stripe formation* can be seen (see Fig. 5.14 middle and bottom left).

Apart from the reduced capacity at the location of a bottleneck, a bottleneck has another negative influence on the crowd. If people are moving towards a bottleneck, sooner or later they will start to interact physically with each other, given that the bottleneck is narrow and people are impatient. To demonstrate this effect, we set up a Monte Carlo simulation, where pedestrians are distributed randomly in front of a bottleneck, with a given crowd density in the interval $\rho \in [0, \rho_{max}]$. ρ_{max} means the maximum density at which simulated pedestrians are standing shoulder to shoulder. To quantify the probability of mutual obstruction, we measure the relative frequency with which two pedestrians overlap, if they are moving a distance towards the bottleneck corresponding to their step size. Figure 5.15 shows the results of the simulation, at a distance 5 m from a bottleneck with zero width, meaning that everyone is moving towards the same focal point.

This helps to understand intermittent [55] outflows and the “faster-is-slower effect” [53]: If a crowd is leaving a room through a narrow door, the flow is smooth if the crowd stays relaxed. However, in an emergency evacuation, the flow can be interrupted sometimes, which reduces the throughput. The reason is that the densities in nervous,

pushy crowds are higher and their distances are smaller, so that mutual obstructions (and clogging due to friction) are more likely.

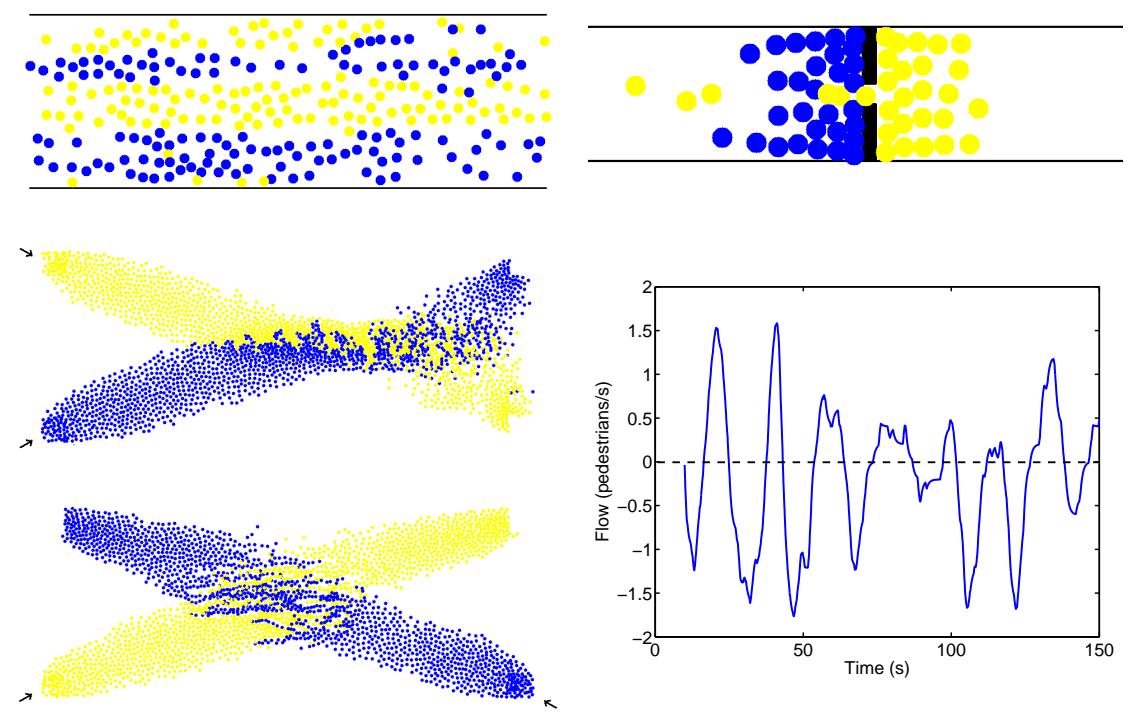


Figure 5.14: Simulation results for the *Elliptical II* specification of the social-force model.

Top left: Lane formation. The color coding is used to distinguish movement in the left direction from movement in the right direction. Middle and bottom left: Illustration of stripe formation, which allows pedestrians to traverse an intersecting pedestrian stream without necessarily having to stop. Top right: Illustration of oscillations in bi-directional flows at a bottleneck. Bottom right: Resulting oscillations in the pedestrian flow, from a computer simulation of the bottleneck. The data were averaged over a 5-second time window.

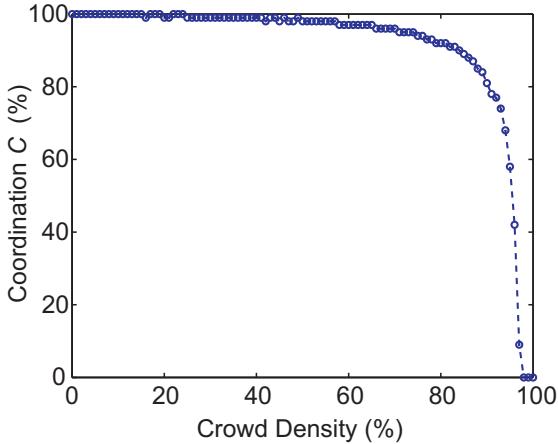


Figure 5.15: Degree of coordination C at a bottleneck. The deviation from 100% reflects the likelihood of mutual obstructions of pedestrians when moving towards a bottleneck. The result is obtained by a Monte Carlo simulation, where pedestrians are distributed randomly in front of a bottleneck, with a given crowd density in the interval $\rho \in [0, \rho_{max}]$. ρ_{max} denotes the maximum density at which simulated pedestrians are standing shoulder to shoulder. The bottleneck has zero size, meaning that everyone is moving towards the same focal point. $1 - C$ is the relative frequency of cases where pedestrians taking one step forward (towards the bottleneck) would overlap. The measurements are made at a distance of 5 meters from the bottleneck.

5.7 Pre-Movement-Time Modeling

For evacuations, not only the movement is of importance, but pre-movement as well. Pre-movement is defined as the time from an alarm or a fire cue until the time when pedestrians actually start to move towards an exit. The traditional way to treat the pre-movement time in pedestrian-evacuation simulations, is to manually introduce pre-specified pre-movement-time distributions as input to the simulation, and then each pedestrian will initially be assigned a random waiting time from the distribution and stall for this time until he or she starts to evacuate.

This treatment of pre-movement time often shows good agreement with empirical data of evacuations, but it uses the assumption that each pedestrian will have a pre-movement time independent of the other evacuees.

We argue here that this assumption is not realistic. Rather it is frequently reported from evacuations and fire drills that the social interaction and spatial configuration of the evacuees has a large influence on the pre-movement time.

For this reason, we will now present data from evacuation experiments on social influence, and work towards a model of pre-movement time, based on social interactions. The assumption behind the model is that pre-movement times are not independent, but rather coupled to the pre-movement times of the fellow persons close by.

In 1999, a number of unannounced evacuations of a cinema in Malmö, Sweden was carried out [14, 103]. The aim of these evacuations was to investigate six different alarm types, and find out how these different alarm types affected the evacuees. In all evacuation experiments, the alarms started after the commercials, but before the movie.

What we will do here is to analyze the data of the experiments with a focus on evacuation simulations, and more specifically on pre-movement-time simulations.

The datasets of the evacuation experiments contain:

- A unique ID of every person and a corresponding seat index, which is organized in a 9×15 matrix (9 rows and 15 columns).
- A distinct pre-movement time for each person, calculated as the time from when the alarm started, until when the person has stood up.
- A group ID for each person, keeping track of which persons came together as a group (and which came alone).

With these datasets one can analyze how the pre-movement time of a person depends on the following factors:

- The person's spatial location.
- The pre-movement times of the people around.
- The pre-movement times of the people within the same group.

5.7.1 Social Influence

To be able to analyze social influence, probability distributions are calculated from the empirical data, which represent the probability that a person is starting his evacuation within the next x seconds, given that a certain fraction of his neighbors has already done so.

As Fig. 5.16 shows, a person reacts strongest if a critical mass of around 40% of his nearest neighbors have already started their evacuation. However, the neighbor sensitivity has a negative slope above this value, which can be explained by a heterogeneous group: Some persons are excited enough to react directly to the action of any of the neighbors, while others do not react at all, or at least very slowly, to the respective neighbors.

If the same procedure is repeated, but this time, instead of neighbors, the members in the same group are used, the characteristics look different (see Fig. 5.16 right). The initial phase is alike, but there is not a strong negative slope, which can be explained with the fact that persons have a strong desire to keep together with the group. The groups have been assigned manually from the video recordings, since one can see which persons arrive to the cinema together, sit close together, talk to each other and so on.

Given that a person is standing up, statistics is compiled to find the probability that this will cause another person in the vicinity to stand up. The surrounding people are classified according to the angle between the two persons, where an angle takes the value 0 in the direction towards the screen, $\pm\pi$ opposite to the screen, $\pi/2$ to the right direction and $-\pi/2$ in the left direction. As Fig. 5.17 shows, the probability is higher of responding to people sitting at rows in front rather than behind oneself. As a result, the pattern in which people are standing up reminds of the propagation of a wave, where people close to the screen are generally standing up first and then people at higher rows follow.

There are two phenomena that seem to balance each other:

1. If people in the front rows react faster, this is more likely to create a wave propagating backwards, since the people in front are visible to everyone.

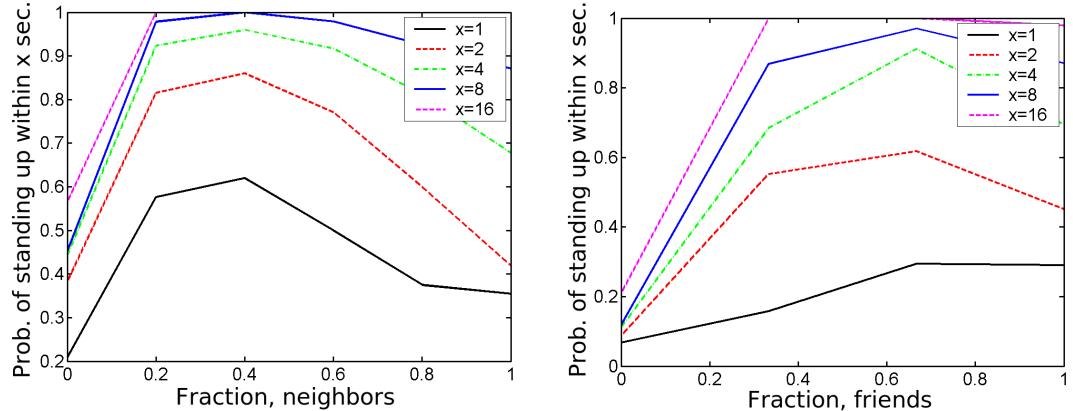


Figure 5.16: Left: Influence of the nearest neighbors on the pre-movement time. Right: Influence of group members (friends). The probability of standing within x seconds is increasing with the fraction of friends who already stood up. The probability to stand up, as a function of the fraction of neighbors (people sitting close by) who stand up looks a bit different. For neighbors, there is a peak probability at intermediate fractions, and then again a decrease for high fractions. This is interpreted as: *If a person has not yet stood up, even though almost all of the neighbors already stood up, the person is likely to remain seated.*

2. People in the front rows have a very limited view on the others and it is, therefore, unlikely that the people in the front will be the first to react.

It turns out that even though waves propagate backwards with higher probability, it is anyway the people on the very front rows which have the longest pre-movement times. This follows from the fact that the front-most people have the least good perception of the rest of the audience, and are therefore less likely to be the first to react to the movement of the others.

5.7.2 Towards a Model of Pre-Movement Time

The analysis of the experiments in the cinema theatre shows that people are influenced by others during the initial phase of a fire evacuation. We will now put together this insight to make a simplistic model. Since the empirical data is only from one set of experiments, the aim is not to make a general model that can grasp all aspects of the

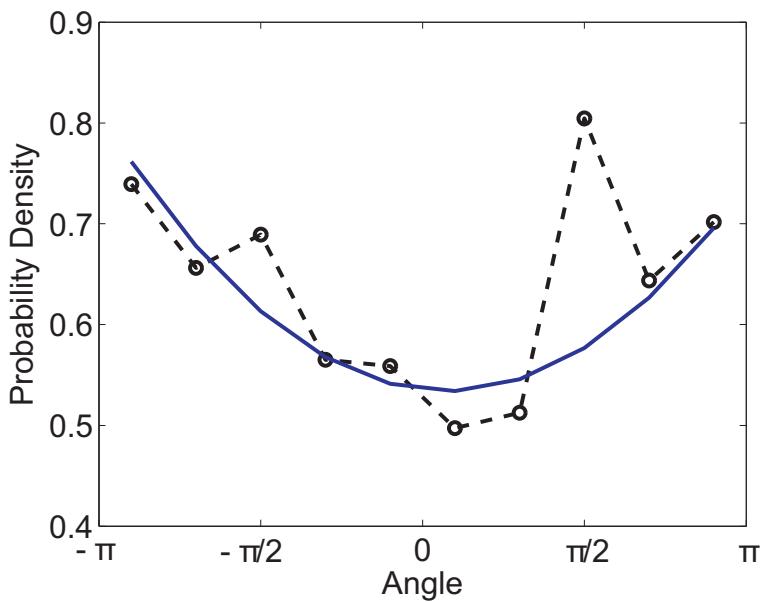


Figure 5.17: Probability that when a person is standing up, another person will follow, as a function of the angle between them. There is a local maximum probability at $\pm\pi$ which is upwards, away from the screen, and there is a minimum probability in the direction 0 which is towards the screen. The reason for the minimum probability towards the screen, is that people are less likely to notice the movements of other people, who are seated in front of them, than of people seated behind them. The solid blue line is added as a guideline to the eye.

pre-movement process during evacuations. Rather, the aim is to present a simple model which can explain the pre-movement process qualitatively.

The aim of performing the simulations is to test the previous conclusions qualitatively, i.e. the model is kept as simple as possible, yet sufficiently sophisticated to give the same characteristics which at least qualitatively corresponds to the evacuation experiments. Testing the conclusions in a quantitative way is not even possible, since then one would need a large number of evacuations with exactly the same alarm type, the same number of participants, etc. Adding all these parameters in the model would just result in an over-fitting of the empirical results, and would not add any value for the understanding of the pre-movement process.

The analysis of the evacuation experiments [14] suggests that the participants were influenced more by their neighbors than by others, which was also assumed in Ref. [35]. This conclusion suggests a model that includes the response of neighbors as an influencing factor. Based on the results of the analysis, the following model is proposed:

$$p_i(t) = \alpha + \beta \frac{n_i(t)}{n_{i,max}} \quad (5.15)$$

where $p_i(t)$ is the probability density that a person gets up from his or her seat at time t , $n_i(t)$ is the number of neighbors who have got up at time t , and $n_{i,max}$ is the total number of neighbors. The index i denotes the seat index in the cinema theatre. In the equation, the total number of neighbors $n_{i,max}$ has a maximum value of eight, according to the 1st order Moore neighborhood.

The parameter α is an individualistic parameter and β describes how much the person is influenced by others.

To test and demonstrate the model, the two cases A2 and B2 from Ref. [14] are simulated. The reason why these two cases are chosen is that these are the two extremes: Case B2 has very clear information in form of an oral message and case A2 has very ambiguous information (an alarm bell). Also, for simplicity, full utilization of the seats in the simulations are assumed.

The time when persons are getting up (behavior 3 in table 3 in Ref. [103]) is simulated, since this is the clearest indicator that a person intends to start his or her evacuation. A homogeneous audience with identical parameters is assumed⁵. The optimal set of parameters $\{\alpha, \beta\}$ are found for the data from the experiments in Ref. [14], by using a *genetic algorithm* [96].

The key parameter is defined as $N(t)/N_{max}$ which corresponds to the normalized number of evacuees that have got up on their feet at time t . A batch of simulations are carried out in the time interval $t \in [0s, 60s]$. As in the experiments [14], seats are organized in a matrix of size of 9×15 .

Two alarm types with different characteristics are chosen, B2 (a female oral message) and then A2 (an alarm clock). Fig. 5.18 shows the simulation results for these two evacuation simulations.

5.7.3 Results

The characteristics of the model is not far from the empirical data (given the simplicity of the model). The asymptotic behavior, where approximately 5% of the crowd takes

⁵ Alternatively, a deterministic model with heterogeneous parameter values could have been used.

an exceptionally long time to react, can not be grasped by the current model, but can be dealt with by assuming a heterogeneous crowd with individual parameters, or even add a group of opposing persons that does the opposite of the others, which was in fact observed in some of the experiments [14].

As expected, the ambiguous case with the alarm clock (case A2 in Ref. [14]) results for a large β parameter and a small α parameter. For the clear cue with the oral female message (case B2 in Ref. [14]), the situation is the opposite, and corresponds to a large α and a small β .

The time where the first person is raising is varying a lot, especially for ambiguous cues. To handle this problem in the simulations, the time when the first person has got up on his/her feet is denoted by $t = 0$. Since there exists only one dataset for each type of alarm, the variance of the starting time is unknown. Therefore it is not intended to reproduce this value with the model. The starting time is expected to vary more, the more ambiguous the cue is, but since there is only one dataset of each alarm type, this issue is left by now.

The simulation results support the claim that social interaction is more important, the higher the level of uncertainty (ambiguous cues). This can be seen as a higher value of the β parameter in the case A2 as compared to the case B2.

The main conclusion of the study is that evacuees are socially influenced by each other, especially during the initial phase of an evacuation. This influence is more pronounced when the information is limited, for example in case of ambiguous fire cues. Results also indicate that the social influence increases with decreasing physical distance. This implies that a person is influenced more by people who are close by, than by people who are further away.

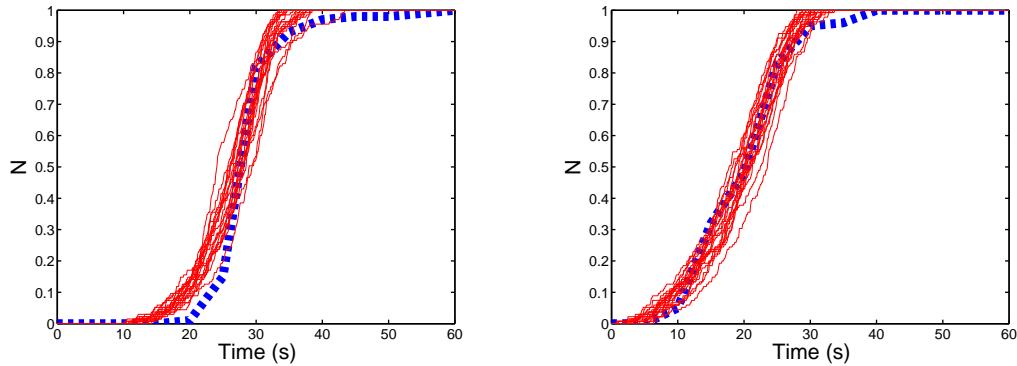


Figure 5.18: Number of persons who have stood up, as a function of time. The thick dashed line corresponds to the empirical data from the experiments in Ref. [14], and the solid lines correspond to 20 individual simulations of each evacuation. Left: Oral message read by a female, with resulting parameters $\alpha=0.22$, $\beta=0.13$. Right: Alarm clock, with parameters $\alpha=0.015$, $\beta=0.23$. The models fit reasonably well to the empirical data, given the simplicity of the model. However, there is a significant deviation at around 15 seconds for the female oral message, which corresponds to the fixed time to listen to the message and process its information. Such cognitive processing time is currently not an (explicit) part of the model. Moreover, the exact starting point, when the first person is standing up, is assumed to vary between different realizations of the same alarm type. Unfortunately, however, only one experiment of each kind of alarm was carried out.

6 Optimization

6.1 Introduction

An interesting application of optimization for pedestrian simulations is to simulate an adaptive/variable scene, where some of the elements can be changed. For these kind of scenarios, one can define an objective function, e.g. *Number of successful evacuees within 60 seconds*, and iterate an optimization loop, adjusting the scene according to the objective function.

One optimization algorithm, which is suitable for this kind of optimization is the *genetic algorithm* [96]. For this algorithm, one defines a population consisting of a number of N *individuals*. Each of the individuals carries a number of parameter values, called *genome*. The details are given in Algorithm 2 in Sec. 5.4.

6.2 Parameter Optimization

To demonstrate the optimization using a genetic algorithm, we chose to optimize a pedestrian intersection (see Fig. 6.1) with improvements as suggested in Ref. [51]. The features of the improved design include separating different flow directions by adding railings, and adding a pillar in the middle to encourage circular traffic. Also, the bottleneck area in the center of the intersection has been enlarged.

120 pedestrians are simulated with the social-force model (see Sec. 5.2). Initially, they are divided in four equally sized groups and inserted at each of the input areas A , B , C and D . The pedestrians are walking towards the center of the intersection and choose to turn left, right or walk straight, with probabilities 1/3 each. When a pedestrian has passed the intersection and has reached the end of the corridor, the pedestrian turns right and reappears at the next road, e.g. when a pedestrian exits at A , he/she will turn to the right and reappear at corridor B .

Now, the aim is to optimize the intersection. We start with a typical design, called the *conventional design* (see Fig. 6.1 left), and specify a number of possible improvements,

referred to as the *improved design* (see Fig. 6.1 right). The objective function is specified as the average flowrate during 240 seconds of simulation. While using a population of 100 individuals in our genetic algorithm, and running an optimization loop, the fitness reaches the values shown in Table 6.1. When the optimization is finished, two simulations are run - one with the conventional design and one with the improved design, using the optimized parameter set. The result (shown in Fig. 6.2) is that the flowrate becomes twice as high, by improving the intersection design.

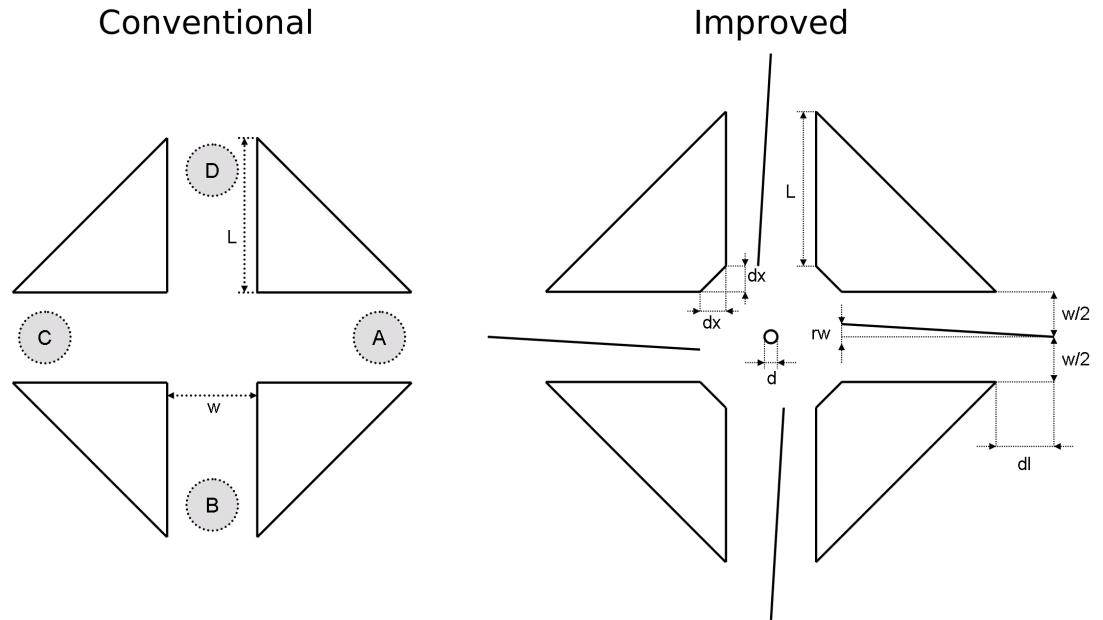


Figure 6.1: The geometry of two different pedestrian intersection designs, which are simulated with the social-force model. Left: Conventional design. Right: Suggested improvement of the conventional design. The values of the parameters d , dx , dl and rw are optimized by a genetic algorithm.

Fixed parameters	Value	Description
N	120	Number of pedestrians (fixed value)
L	7 m	Length of each corridor (fixed value)
w	2.5 m	Width of each corridor (fixed value)
Variable parameters	Value	Description
rw	0.6 m	Transversal displacement of the endpoint of the railing
d	0.0 m	Diameter of the pillar
dx	0.7 m	Enlargement of each side of the central area
dl	1.5 m	Longitudinal displacement of the endpoint of the railing

Table 6.1: Optimal parameter values resulting from our evolutionary parameter optimization. See Fig. 6.1 for an illustration of the parameters.

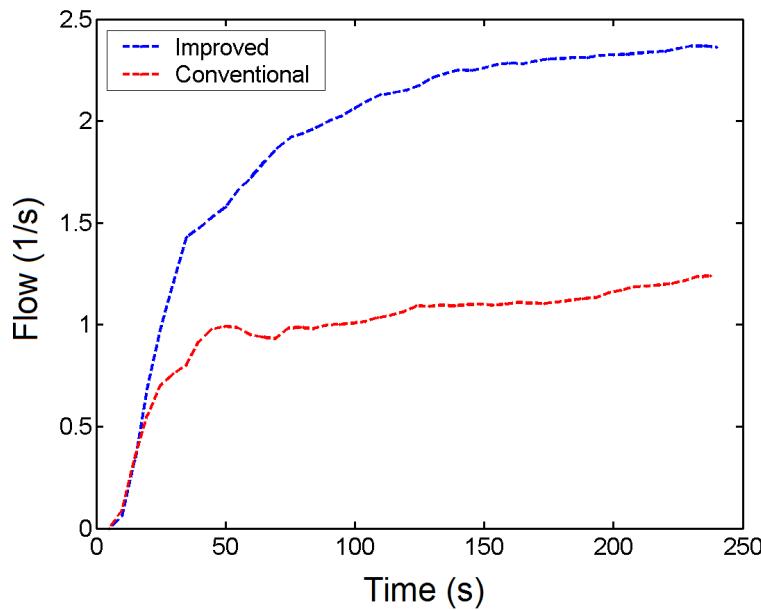


Figure 6.2: Flowrates as functions of simulated time for the conventional intersection compared to the improved intersection with optimized parameters (see Table 6.1). Note that, by making the improvements of the design, the throughput may reach twice as high values as compared to the conventional design.

6.3 Flow Optimization by Boolean Grids

When a large group of pedestrians is evacuated from a building, one of the most dangerous locations is at the doors. In emergency situations when pedestrians panic and fear for their lives, they tend to force their way out, even if the exits are jammed, which is resulting in clogging in front of the exits. This is much more pronounced during evacuations than under normal conditions.

It has been pointed out that it is possible to increase the outflow by suitably placing a pillar or some other type of suitable obstacle in front of the exit. This can reduce the pressure in front of the door, thereby decreasing the magnitude of clogging and improving the overall flow [53, 51].

We have seen that a genetic algorithm can successfully optimize the dimensions of a pre-defined design with the help of an objective function, e.g. maximization of the flow. We will now go one step further and optimize a structure with a minimum amount of pre-specification of how this structure should look like.

To investigate how the architectural infrastructure in the vicinity of a door, or other bottleneck, shall be constructed to maximize the pedestrian outflow under evacuation conditions, we present a method [78] based on a genetic algorithm [96].

6.3.1 Boolean-Grid Representation

In previous optimization studies [21, 50, 34], alternative designs have been compared and parameters have been optimized for certain pre-specified designs, but the topologies of the geometrical designs were always given. To allow for more freedom in the evolution of the infrastructure, we choose a representation similar to the one in Ref. [27], which uses a Boolean grid. In this grid, 0 means *no obstacle* and 1 means *an elementary obstacle*, similar to cellular automata models. These grids are denoted as $X_{i,j}^\alpha$, where $\alpha = 1..N$ is the index of the grid, $i = 1..m$ the row, and $j = 1..n$ the column in the grid. To create an initial “population” of obstacle grids $X_{i,j}^\alpha$, we use the following scheme:

1. Initialization rule:

$$\text{Set } X_{i,j}^\alpha = \begin{cases} 1 & \text{with probability 0.5} \\ 0 & \text{otherwise} \end{cases}, \forall \alpha, i, j. \quad (6.1)$$

2. Update rule (clustering):

Update all $m \times n$ cells in parallel, c number of times, where c is a clustering constant. The clustering is made with the following non-linear updating rule, where the constant d denotes the size of the neighborhood taken into consideration (a Moore neighborhood of degree d):

$$s_{i,j} = \sum_{k,l} X_{k,l}, \text{ for } k \in [i-d, i+d] \text{ and } l \in [j-d, j+d] \quad (6.2)$$

$$p_1 = \frac{1}{2} + \frac{\arctan(k[2s_{i,j} - (2d + 1)^2])}{\pi}. \quad (6.3)$$

$$\text{Set } X_{i,j}^\alpha = \begin{cases} 1 & \text{with probability } p_1 \\ 0 & \text{otherwise} \end{cases} \quad (6.4)$$

Equation (6.2) determines the number of cells containing elementary obstacles, within the Moore neighborhood of degree d , i.e. the $(2d + 1)^2$ closest cells, around the cell at row i and column j . Then, Eqs. (6.3) and (6.4) are assigning an elementary obstacle for cell i, j , with a probability depending on $s_{i,j}$. The constant k determines the (inverse) level of stochasticity in the clustering. In the limit $k \rightarrow \infty$, the clustering rule becomes deterministic, i.e. each cell is updated according to the majority value in the local neighborhood.

The use of the clustering step is to create smooth clusters of solid non-fragmented obstacles from the scattered obstacle grids that are created by the *initialization rule*.

Table 6.2 contains the default values of the different parameters, but let us first vary each of the parameters, to illustrate their respective influence on the resulting grids. Starting with the clustering constant c , Fig. 6.3 shows a grid created by the *initialization rule*, and the same grid after a number of applications of the clustering. After only a few iterations, a structure has formed from the initial random grid. This structure is stable in the sense that it does not change significantly between iteration 5 and iteration 100.

Next, we vary the neighborhood size d . As can be seen in Fig. 6.4, the structure is more detailed the smaller the value of d is. Therefore, this parameter can be used to set the preferability of either detailed and complex grids (small d) or simpler grids (large d).

Finally, five different values of k are used to illustrate how k influences the clustering. Since k determines the level of stochasticity, it can be seen in Fig. 6.5, that a low value

of k creates a “noisy” grid with small isolated clusters of elementary obstacles, and a large value of k creates a grid with less isolated small clusters of obstacles.

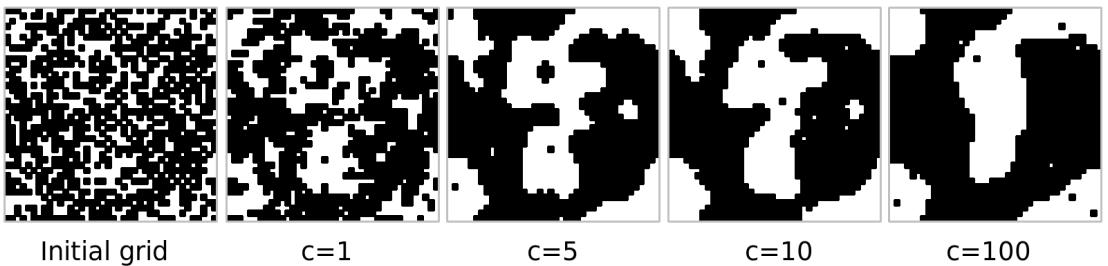


Figure 6.3: An illustration of the effect of the clustering rule, for different values of c , i.e. different number of times to sequentially apply the clustering. The other parameter values are given in Tab. 6.2.

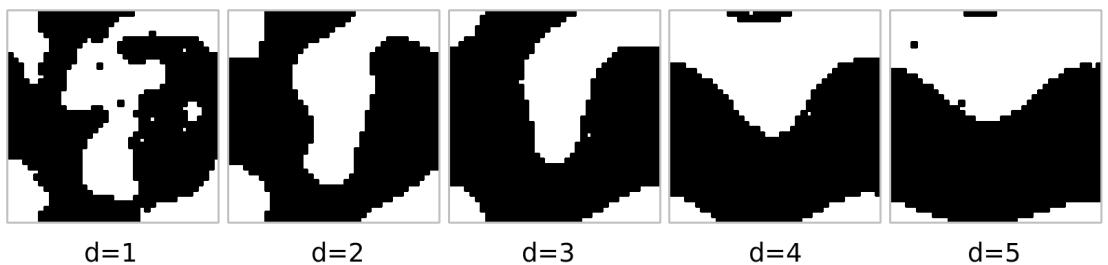


Figure 6.4: An illustration of the effect of applying the clustering rule 10 times, for different neighborhood sizes d . Note that the level of detail is decreasing by increasing neighborhood sizes d . Therefore, one can use this parameter to control the complexity of the resulting obstacle grid. The other parameter values are given in Tab. 6.2.

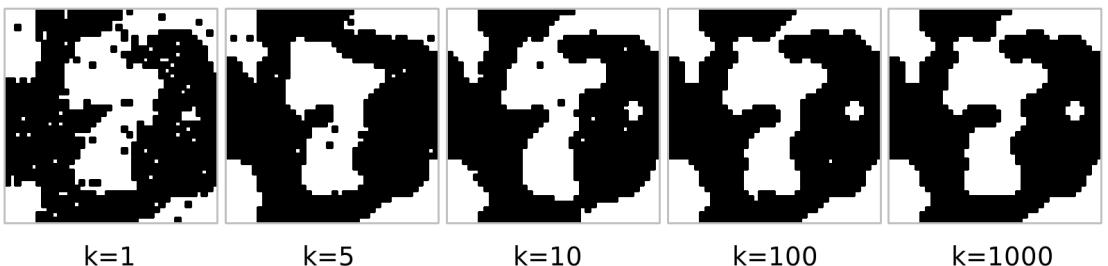


Figure 6.5: The effect of the level of stochasticity in the clustering rule. Note that larger values of k are better for removing noise (small islands), which results in smoother obstacles. The other parameter values are given in Tab. 6.2.

The advantage of generating these obstacle grids rather than using conventional design elements such as pillars, walls, and railings is that a number of objects with variable sizes and shapes can be created, without having to pre-specify the number and shape of these objects.

If properly designed and located, such “obstacles” can surprisingly improve the throughput, efficiency and safety of pedestrian flows, while reducing the travel and evacuation times.

To reach this, we have implemented a genetic algorithm [78, 96] to improve the clustered grids for evacuation scenarios, according to the following scheme:

1. Generate a population of N obstacle matrices $X_{i,j}^\alpha$, where $\alpha = 1..N$, $i = 1..m$, $j = 1..n$. The generation is made by the *initialization* and *update* rules, defined in Eqs. (6.1) to (6.4).
2. Choose four geometrical setups (grids) at random and make a simulation with the social-force model (see Sec. 5.2), running them during 3,600 simulated seconds each.
3. Assign a fitness value to each of the four geometrical setups, e.g. according to the average outflow or traveling time of the pedestrians. Keep the two setups with the best fitness, $X_{i,j}^\alpha$ and $X_{i,j}^\beta$, and update the two worst ones, for all i and j , according to the following rule with random numbers $p_\alpha = \text{rnd}[0, 1]$ equally distributed between 0 and 1:

$$X_{i,j} = \begin{cases} X_{i,j}^\alpha & \text{with probability } p_\alpha \\ X_{i,j}^\beta & \text{otherwise.} \end{cases}, \forall i, j. \quad (6.5)$$

Then smoothen the outcome by the *update rule* described in Eqs. (6.2) to (6.4).

To get an idea what can be improved by adding the obstacles, we run a series of simulations of a $20\text{m} \times 10\text{m}$ corridor, which is terminated by a 0.55 m wide door in the middle of one end of the corridor (see Fig. 6.6). The simulations are made with the *Elliptical II* specification of the social-force model (see Sec. 5.2) with the parameters $A = 0.1$, $B = 2.8$, $\tau = 0.8$, $\Delta T = 2$, $\lambda = 0.1$, and $N = 8$ (number of people to react to).

The noise term $\xi(t)$ is specified as a force vector with a random direction on the interval $0..2\pi$ and a random magnitude on the interval 0 to 0.5. The angle and magnitude of this force are drawn from a uniform distribution, two times, every simulated second.

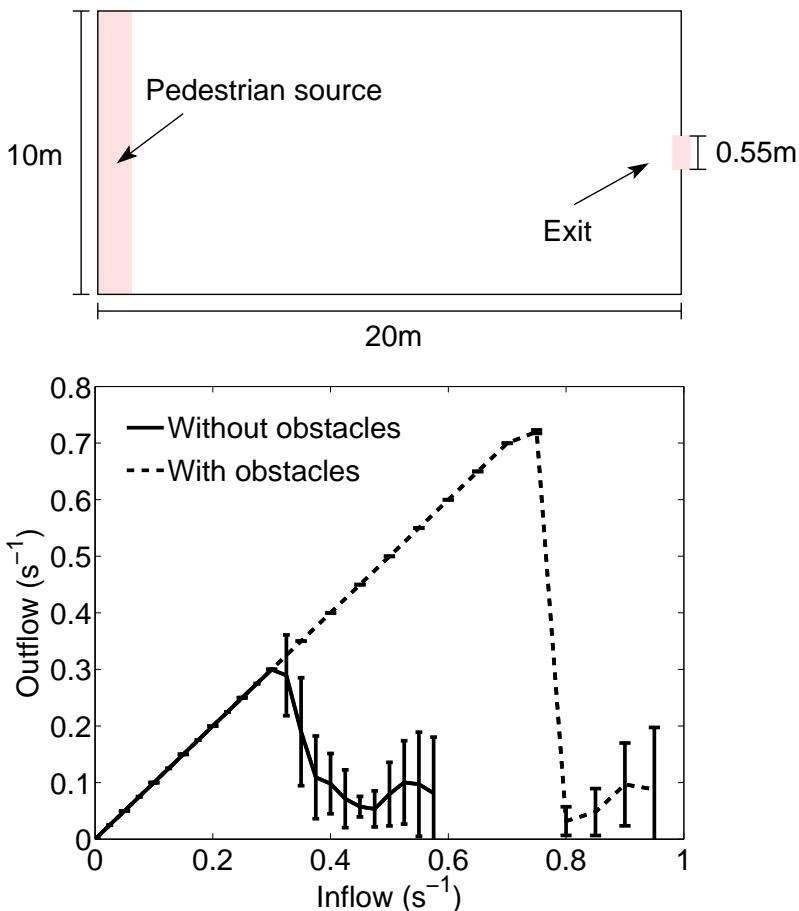


Figure 6.6: The result of a series of simulations of a $20\text{m} \times 10\text{m}$ corridor, terminated by a 0.55 m wide door in the middle of one end of the corridor. Top: The specification of the simulation scenario. Pedestrians are introduced at the left edge of the corridor and are leaving the corridor through the door at the right side. Bottom: The results of a series of simulations for varying inflows. The solid line corresponds to the original scenario without any obstacles, and the dashed line corresponds to the same corridor, but with obstacles added in front of the door, resulting from the optimization process. Each data point in the graph corresponds to the mean value and the error bars to the standard deviation for 5 simulations, 86,400 seconds (24 hours simulated time) each. Note that the graph is initially linear, but above a certain point (the capacity), a jam is building up in front of the door, which makes the throughput lower (because of incoordination and blockages in the crowd). For computational feasibility, the number of pedestrians in the simulation is limited to 100. Without such limitation, the jam is expected to grow infinitely, when the inflow is above the capacity, which would give an even sharper transition.

Parameter	Meaning	Default value
c	Number of times to sequentially apply the clustering rule.	10
d	Neighborhood size used by the clustering rule, i.e. a Moore neighborhood of degree d is used ($(2d + 1)^2$ cells).	1
n	Number of rows in the grid.	40
m	Number of columns in the grid.	40
k	A constant, defining the level of stochasticity in Eq. (6.3) of the clustering rule.	10
N	Number of grids to generate, i.e. the “population” used by the genetic algorithm.	100

Table 6.2: A summary of the parameters and their meanings. Their values are set according to their default values, unless different values are explicitly stated in the text.

Additionally, there is a second (circular) inter-pedestrian force which reflects the physical repulsion. This force is specified by Eq. 5.5 and parameters $A = 3$, $B = 0.2$, and $R = 0.15$, together with Eq. 5.4 and $\lambda = 0.1$.

A series of simulations are carried out, while varying the inflow of pedestrians inserted into the corridor. At the same time, we measure the outflow of people through the door. The simulations reveal that two phases exist. For inflows lower than the capacity of the door, the outflow corresponds exactly to the inflow. For inflows larger than the capacity of the door, on the other hand, a jam is forming in front of the door, since the pedestrians are not able to leave the corridor at the same pace as they are introduced into the corridor. In this second phase, when there is a jam in front of the door, there is a drop in the outflow, because of mutual obstructions and incoordination in front of the door.

Now, we run the same series of simulations, but with an obstacle grid added in front of the exit. The obstacle grid we choose, is the one with the best fitness (lowest average traveling time) from the optimization with the genetic algorithm. The new series of simulations, with the obstacle grid added in front of the exit, shows that the capacity has been increased significantly by putting the obstacle grid in front of the exit. In Fig. 6.6, the two different scenarios (with and without obstacle grid) are compared. By adding the optimized obstacle grid in front of the exit, the capacity of the exit more

than doubles. However, there is still a transition at the capacity point, which makes the outflow drop significantly. It is clear that the capacity can only be increased up to a certain maximum limit, but is there no way to prevent the outflow from dropping significantly, for inflows above the capacity? This would certainly be a good property of the design. In order to investigate this, we run a new optimization with an inflow $\rightarrow \infty$ and an upper limit of 300 pedestrians. In other words, there will always be 300 pedestrians present in the simulation, and the aim of the genetic algorithm is now to find an obstacle grid that will prevent the exit from jamming, and keep the outflow as high as possible, even when there is a crowd of 200 pedestrians waiting in front of the exit.

Due to the non-linear nature of the interactions, it is apparently better to create a smaller bottleneck before a bigger one, so that the transition between the parts with different capacity is not that sharp. It becomes obvious that the weakest link does not set a strict upper limit for the flow, but rather the weakest link may be improved by changing the infrastructure preceding it.

6.3.2 Results

From the various resulting improved solutions, we conclude that the following design elements are potentially favorable (see Fig. 6.7):

1. Asymmetry, as the evolving obstacles at the left side are usually different from the obstacles on the right side.
2. Funnel shape, i.e. an eventual decrease in the walkable width towards the bottleneck.
3. Zig-zag shape, i.e. a series of turns in different directions.
4. Compartment structure similar to the funnel shape, but with an increase in the walkable width after one bottleneck before another, which creates an area of lower density compared to the rest of the crowd.

As an example, an asymmetrically placed pillar in front of an exit contains the favorable elements 1 and 2 in the above list. A particular zig-zag-shaped design has been recently suggested to improve the safety of egress routes in sports stadia [51].

It has been demonstrated how flow optimization can be reached by adding obstacles without having to pre-define the shape and number of obstacles. However, it can never

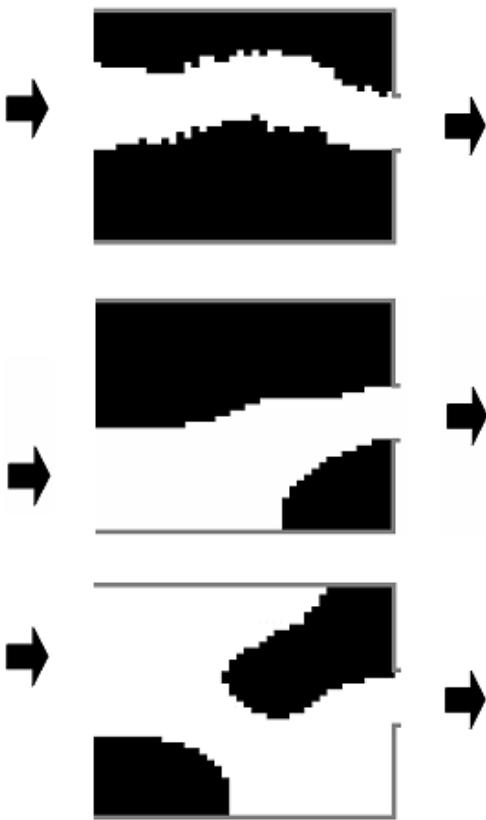


Figure 6.7: Examples of obstacle grids resulting from the Boolean-grid optimization. Top: Zig-zag shape. Middle: Funnel shape. Bottom: Compartment shape. Note that the top part may be dangerous during evacuations since people may get stuck there. However, such designs can not always be avoided during the automatic optimization, since if the flow is below the bottleneck capacity, nobody gets stuck there. Therefore, even though automatically optimized geometries may be very useful, they should never be implemented directly without first being evaluated by experts.

be guaranteed that the solutions may not have potentially negative properties which do not show up during their automatic evaluation. Therefore, they should never be implemented directly without first being evaluated by experts.

7 Hajj, the Pilgrimage to Makkah

7.1 Introduction

The Muslim pilgrimage, called the *Hajj* [119, 98], is an annual event performed in the sacred areas of Makkah, Mina, Muzdalifa, and Arafat, in the Kingdom of Saudi Arabia. The event is known to be particularly crowded, with millions of pilgrims performing their religious duties with strict constraints both regarding time and space. Hajj has a history of many large crowd disasters [125]. As the number of pilgrims has been increasing every year, this has put more pressure on the different authorities organizing the event. In recent years, more effort has been spent to approach this problem by scientific means, in terms of crowd simulation [125, 77, 8, 85], grouping and scheduling of pilgrims [4], crowd management [42, 2], organization of the streets [5, 6], luggage management [7], video evaluation [3] and constructural changes [1].

Since the Islamic calendar is using Hijri years, rather than the years used in the Gregorian calendar in Western countries, the post-fix “H” is used to denote Hijri years and “G” is used for the Gregorian years. By this denotation, the year 2006G corresponds roughly to the year 1426H. The Hajj is performed during the month Dhu al-Hijjah, which is the 12th and last month of the Islamic calendar.

7.2 Validation of the Video-Analysis Software

Before the Hajj in 1426H (January 2006G), one dozen cameras were mounted around the tent city Mina, next to Makkah, in the Kingdom of Saudi Arabia, in order to gain the data required for a better understanding and preparation of the Hajj. From all these cameras, videos were recorded 24 hours a day, during three days, from 10th to 12th of January 2006.

In early 2006, a project was initiated by the Ministry of Municipal and Rural Affairs, in the Kingdom of Saudi Arabia with the aim of evaluating video material in a scientific context. The first stage of the project was to evaluate the video-tracking software (see

Chapter 2) on real-life crowd videos over 24 hours, which includes a large variability both in crowd density and in light conditions.

Since each video camera is different (different lens, different light intensity, different contrast, etc.) a calibration has to be made separately for each camera. This calibration is based on comparing 30 reference frames, where the pedestrian' heads were counted manually. With this information, the error between automatic counting and manual counting can be determined, and then the optimal set of parameters (minimizing the error) can be obtained.

The first test is made to check that the software can operate in the full spectrum of crowd density, ranging from 0 pilgrims to exceptionally high crowd density. This is made possible by neglecting the temporal dimension of the videos and create 30 snapshots from each video, during different level of crowd density and different time of the day (which gives changing light conditions, see Sec. 7.3).

A rectangle with arbitrary position and size is put onto each snapshot, and the number of pilgrims inside the rectangle is then determined both by the software (Fig. 7.1 left) as well as by two persons (Fig. 7.1 right). Using 15 of the snapshots, the parameters (see Tab. 2.2 in Sec. 2.3.2) for the video software are calibrated by Algorithm 1 in Sec. 2.3.6. The calibrated software is then applied on the remaining 15 snapshots in order to evaluate its ability to estimate the density in an accurate way.



Figure 7.1: The green box, used for manual head counting. On the computer, a marker is put on each head to prevent double counts or missing counts. This method was used in 2006, but has later been improved (see Sec. 2.3.6).

Next, when the ability of the video-analysis software to accurately determine crowd density has been made, the next step is to check its ability to analyze the temporal dimension of the videos as well, i.e. the ability of the software to extract not only the density but also the flow and the average velocity of pilgrims. To be able to do this, however, a substantially higher effort is needed in the manual evaluation of the videos, used as a reference for flow and velocity. See Appendix C for technical details.

Since the main interest is *not* explicitly on how individual pilgrims behave, but rather on how they move as a crowd, the *macroscopic* approach to video analysis from Sec. 2.3 is used to extract aggregated quantities of the pilgrim crowds. To assess the accuracy of the results from the video-tracking software, manually obtained data are required. The scale of the data (hundreds of thousands of pilgrims within 24 hours *per camera*) does not make it possible to manually count *all* of the pilgrims. Rather, we sample the data by the following scheme:

- Playing the videos on a large screen TV (see Fig. C.3), 10 times slower than real-time,
- counting the number of pilgrims who pass a certain cross-section line, within 5 seconds, and
- repeating the procedure every 10 minutes.

This scheme is then carried out for the full 24 hours, and the same counts are carried out by two different persons to get a certain evaluation of the human counts.

We separate the counting into two different directions, i.e. pilgrims who pass the cross-section line in one direction we call the *in* direction, and the perpendicular direction we refer to as the *out* direction

The Jamarat Bridge (see Fig. 7.2) is one of the main focal points during the Hajj. Therefore, the flow measurements are separated into flows towards the Jamarat Bridge (the *in* direction), and back from the Jamarat Bridge (the *out* direction).

As a result of the manual counting, there are 576 measurements ($24 \times 6 \times 2 \times 2$) for every 24-hours of video recordings (6 counts per hour, two counting persons and two walking directions.). These measurements give enough information, both for comparing the flow rate into the *in* (*out*) $Q_i(t)$ ($Q_o(t)$) direction, as well as the cumulative flows in the *in* (*out*) $N_i(t)$ ($N_o(t)$) direction. With cumulative flow, we mean the total number of pilgrims who pass the cross-section line from a reference starting time t_0 until time t .

Within the time window $t_0..t_0+5$ seconds, n_o pilgrims pass the cross-section line in the *in* direction and n_o pilgrims pass the line in the *out* direction. From this information,

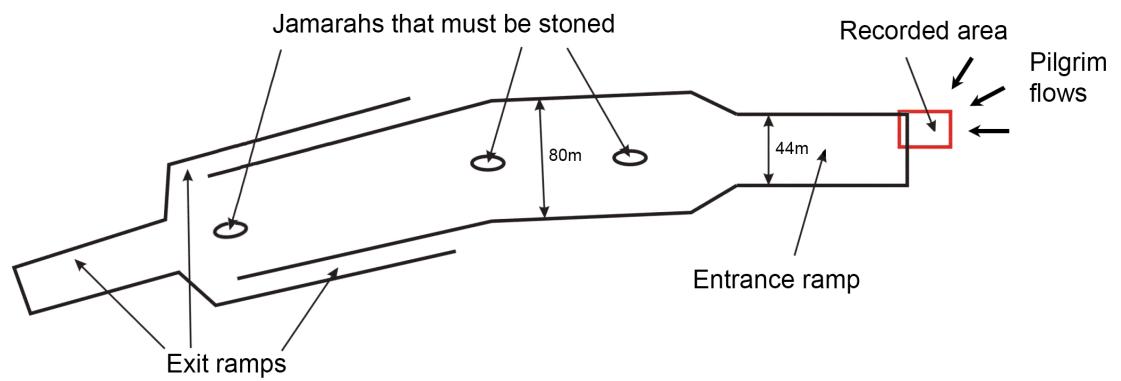


Figure 7.2: Top: Illustration of the previous Jamarat Bridge. The pilgrims have to throw seven pebbles against each of the three Jamarahs (pillars/walls) in the stoning-of-the-devil ritual. The red rectangle marks the recorded area where video recordings were made. Bottom: One of the elliptically shaped Jamarahs.

we estimate the flow in the in direction as $Q_i(t) = n_i/5$ pilgrims/second (and vice versa for the other direction). Since the counting is performed every 10 minutes (600 seconds), the cumulative flow (total number of pilgrims since time t_0) can be estimated by $N(t) = \int_{t_0}^t dt' Q(t') \approx 600 \sum_{i=0}^{(t-t_0)/600} Q(t_0 + 600i)$.

Now, with the time series of the flow, an automatic counting procedure is made for the same videos with the calibrated video-analysis software. The comparison is shown in Fig. 7.3.

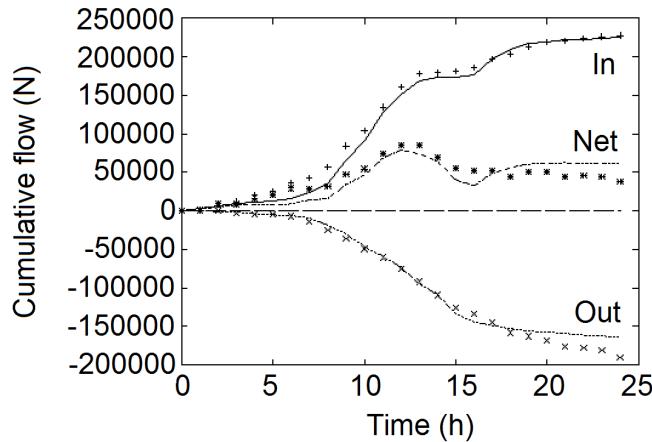


Figure 7.3: Example of the cumulative flows during the 24 hours of the 10th day of the Hajj in 1426H, for one street in Mina, Kingdom of Saudi Arabia. The lines correspond to the automatic counting and the markers correspond to the manual counting.

Note that the dataset used for calibrating and validating the software (30 snapshots) is different from the dataset used for validating the accuracy of the software (5-second counts, repeated every 10 minutes).

The method of counting the flows manually, during 5 seconds, and repeating the procedure every 10 minutes is primarily done to assess the results of the video-tracking software. Even though the manual counting data can be used to estimate the total number of pilgrims during 24 hours, to get more reliable results one should run the video-tracking software with a sampling frequency much smaller than 1 minute, to minimize sampling error.

In order to evaluate how the error depends on sampling frequency, high resolution time series are needed. Since it is not feasible to manually count every single pedestrian during 24 hours, we make the assumption that the variance in the time-series obtained from the video-analysis software corresponds to the variance in the real data. Now, we use the time series from 5 different cameras and sample them with different frequencies, from 1 second up to 60 minutes, and measure the absolute value of the relative error

$|E|$, as defined in Eq. (2.6). Figure 7.4 shows the results of the comparison, where the mean error is displayed as a solid line and the standard deviation as dashed lines. With this method, the error due to the 10-minute sampling is estimated to be slightly below 10%.

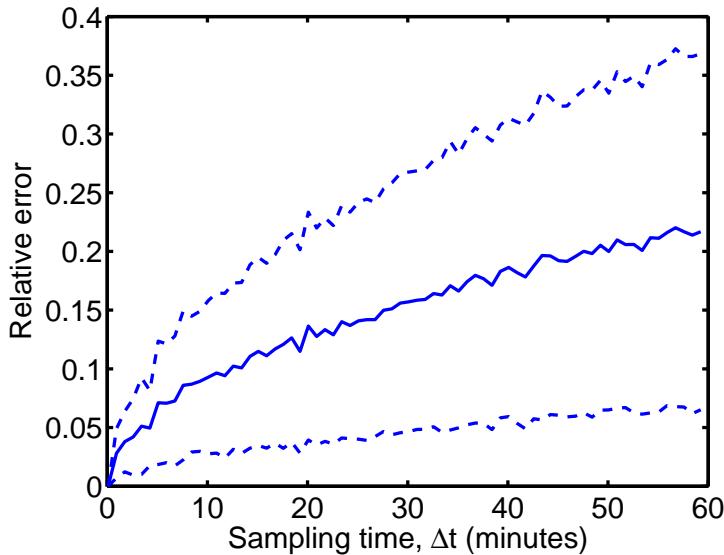


Figure 7.4: An investigation is made on how the remaining error after 24 hours depends on the sampling frequency. The mean error is displayed as a solid line and the standard deviation as dashed lines.

7.3 Light Dependence

Since we are interested in real-life conditions, the changing light conditions during daytime and nighttime need to be taken into account and compensated for. The situation becomes even more complicated since some video cameras automatically switch to black-and-white recording during nighttime, which causes the brightness and the dynamics of the video to look very different between day and night (see Fig. 7.5).

To handle this problem, we use two separate parameter sets, one for daytime and one for nighttime. A comparison of the errors using the same parameter set compared to using two different parameter sets is shown in Fig. 7.5. Both curves have some error fluctuations, but when using only one parameter set, there is a clear *systematic* error due to the changes in light between daytime and nighttime. The error is defined in

Eq. (2.6), and the peaks occur at video frames when the number of people is low, i.e. crowd conditions are not dangerous (when the denominator Q_m is very low, the relative error easily becomes large).

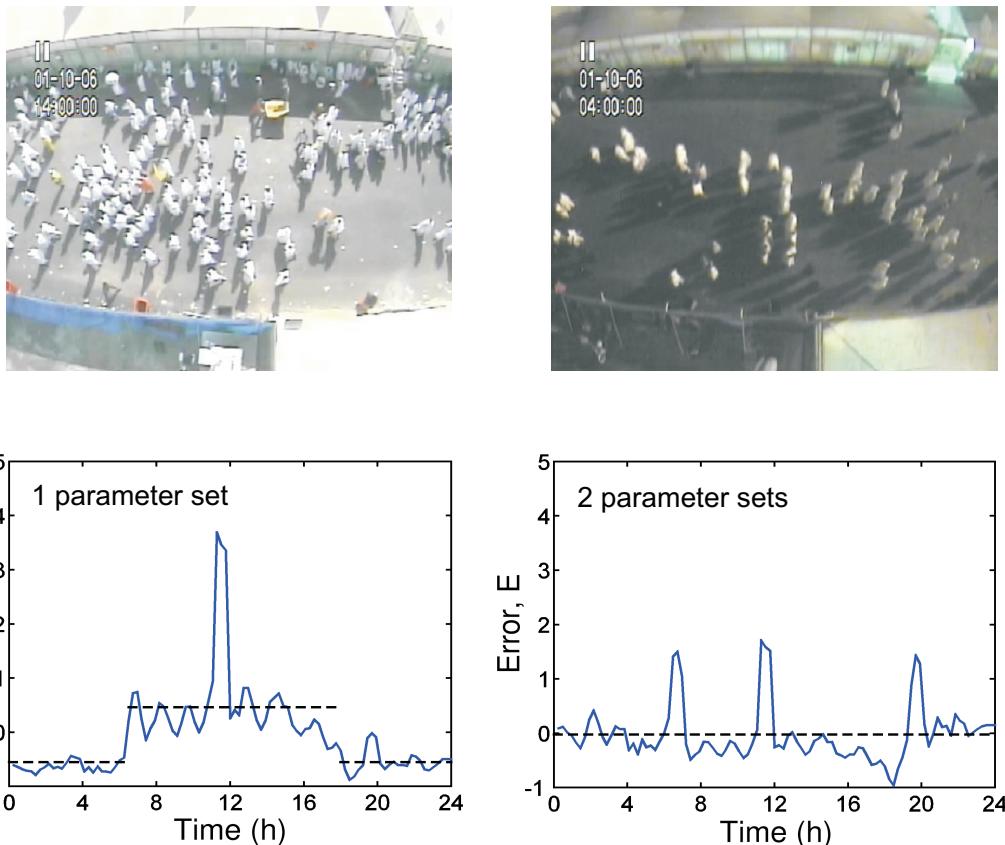


Figure 7.5: Snapshots from video recordings during day (top left) and night (top right).
 Bottom left: The relative error, defined in Eq. (2.6), as a function of the time of the day, with homogeneous parameters (only one parameter set).
 Bottom right: Relative error for two different sets of parameters, one for daytime and one for nighttime. Note that the *systematic* error vanishes when two parameter sets are used instead of only one parameter set. The error peaks occur at video frames when the number of people is low (when the denominator Q_m is very low, the relative error easily becomes large)

7.4 Analysis of a Crowd Disaster

Over the years, the Hajj, in Mina, Kingdom of Saudi Arabia, has seen many tragic crowd disasters and therefore, surveillance cameras were mounted at various places that were particularly crowded during previous pilgrimages.

Then, during the Hajj in 1426H, at the 12th of Dhu al-Hijjah (12th of January 2006), another crowd disaster occurred. This crowd disaster occurred before the 44-meter wide on-ramp to the (previous) Jamarat Bridge, and led to more than 360 fatalities.

One of the video cameras happened to be located directly above the area where this crowd disaster occurred (see Fig. 7.2). This made it possible to make scientific evaluation of the crowd dynamics which eventually caused the disaster.

The videos were made available for scientific analysis with the purpose to come up with strategies minimizing the risk of similar crowd disasters in the future (see Sec. 7.5).

The corresponding video camera was mounted on a 35-meter high pole, pointing perpendicular to the ground and covering an area of $28\text{ m} \times 23\text{ m}$, which corresponds to half of the width of the entrance ramp to the previous Jamarat Bridge.

Snapshots from the video recording (see Fig. 7.6) show how the density is building up. Interestingly, even for very high average density $\varrho = 7$ pilgrims per m^2 , there is a large variation in local density ρ . Small empty areas can even be seen.

By measuring the flow of pilgrims (see Fig. 7.7 left) we see that, within 15 minutes only, the flow breaks down from a level which is close to capacity (i.e. maximum flow) to sometimes very small values. However, the breakdown happens already half an hour before the crowd disaster.

A closer analysis shows that, after the density exceeded capacity (when the flow started to decrease), the crowd entered a very critical phase: The *decrease* in flow resulted in a steady *increase* of density, which further *decreased* the flow, and so on. Therefore it would have been crucial at this time (and preferably even before) to reduce the inflow considerably to prevent an increase of the crowd density.

By integrating the flow over time, we obtain the total number of pilgrims who have entered the Jamarat Bridge. The slope of this curve gives us the smoothed (averaged) flow (see Fig. 7.7 right).

Apparently there was a considerable decrease of the flow before and during the disaster, but the flow never vanished over a significant period of time, even though the flow experienced a transition from a steady (laminar) flow to stop-and-go flow (see Sec. 3.6) at around 11:53.

The stop-and-go waves persisted for more than 20 minutes. While the crowd density continued to increase, a second transition took place at 12:19, when the crowd turned into a turbulent phase (see Sec. 3.7).

About 10 minutes after the crowd turned turbulent, the disaster started.

At the front edge of turbulent waves, strong forces act onto the people, which give them a push. This is especially dangerous for the persons next to boundaries, e.g. walls, since these persons may be pushed into the boundary and might get injured. People may also fall, and in view of the extreme crowd density, it is difficult to get up again. It has been measured that the forces in a crowd can reach magnitudes up to 4,500 Newton per meter [122].

For practical reasons, it is not enough to be able to classify stop-and-go waves and crowd turbulence once they occur. It would be highly desirable to be able to anticipate these phenomena before they occur (which could potentially anticipate the disaster as well). If such an anticipation would be possible, one could counteract the development of a potentially critical situation.

The question is now: *Which quantity will best anticipate the time and location where critical situations occur?*

Let us start by looking at the crowd density and how it evolves both as a function of time and as a function of space. The crowd density increases monotonically until the time of the disaster (Fig. 7.8 a). Therefore, the density can explain the time at which the disaster occurs. However, while looking at the density as a function of space (Fig. 7.8 b) the area with the highest density does not match with the area of the disaster. The problem with using density as a quantity for anticipating the level of criticality is that the density can not reflect the movement of people (especially the dangerous irregular movement).

Therefore, in an attempt to grasp the irregular turbulent motion patterns, we apply the Curl operator on the velocity $\text{curl}(\vec{V}) = \vec{\nabla} \times \vec{V}$. Unfortunately, the Curl operator (see Fig. 7.8 c and d) performs even worse than the density, to assess the level of criticality of the crowd both as a function of time and space.

Next, we evaluate the negative velocity divergence $-\text{div}(V) = -\nabla \cdot \vec{V}$ as a quantity to measure the compression of the crowd. Unfortunately, also this quantity does not manage to grasp the level of criticality of the crowd (see Fig. 7.8 e and f).

Finally, a quantity which can combine the good performance of density as a function of time with the irregular properties of crowd turbulence is proposed by the *crowd pressure*, which is defined as: $P(\vec{r}, t) = \rho(\vec{r}, t) \text{Var}_{\vec{r}, t}(\vec{V})$. Note that this is not Newtonian pressure,

but rather a gas-kinetic pressure. High pressure is reached if both the crowd density as well as the velocity variance within the crowd are high. When evaluating the crowd pressure from the data, we see that it has predictive power, both as a function of time and space. By averaging the pressure over space, we see that the pressure is increasing during 30 minutes before the disaster occurs. By averaging the pressure over time, before the disaster occurred, one can see that the area with the highest pressure corresponds to the area where the disaster occurred. Figure 7.8 (g and h) shows the crowd pressure. Therefore, from the evaluated quantities, the crowd pressure is the only quantity which seems to anticipate the disaster both in time and in space.

Note that, since the number of crowd disasters which have been scientifically evaluated in the past are very limited, one can not be certain of the generality of these measurement quantities and description of how the dynamics of the crowd evolves.

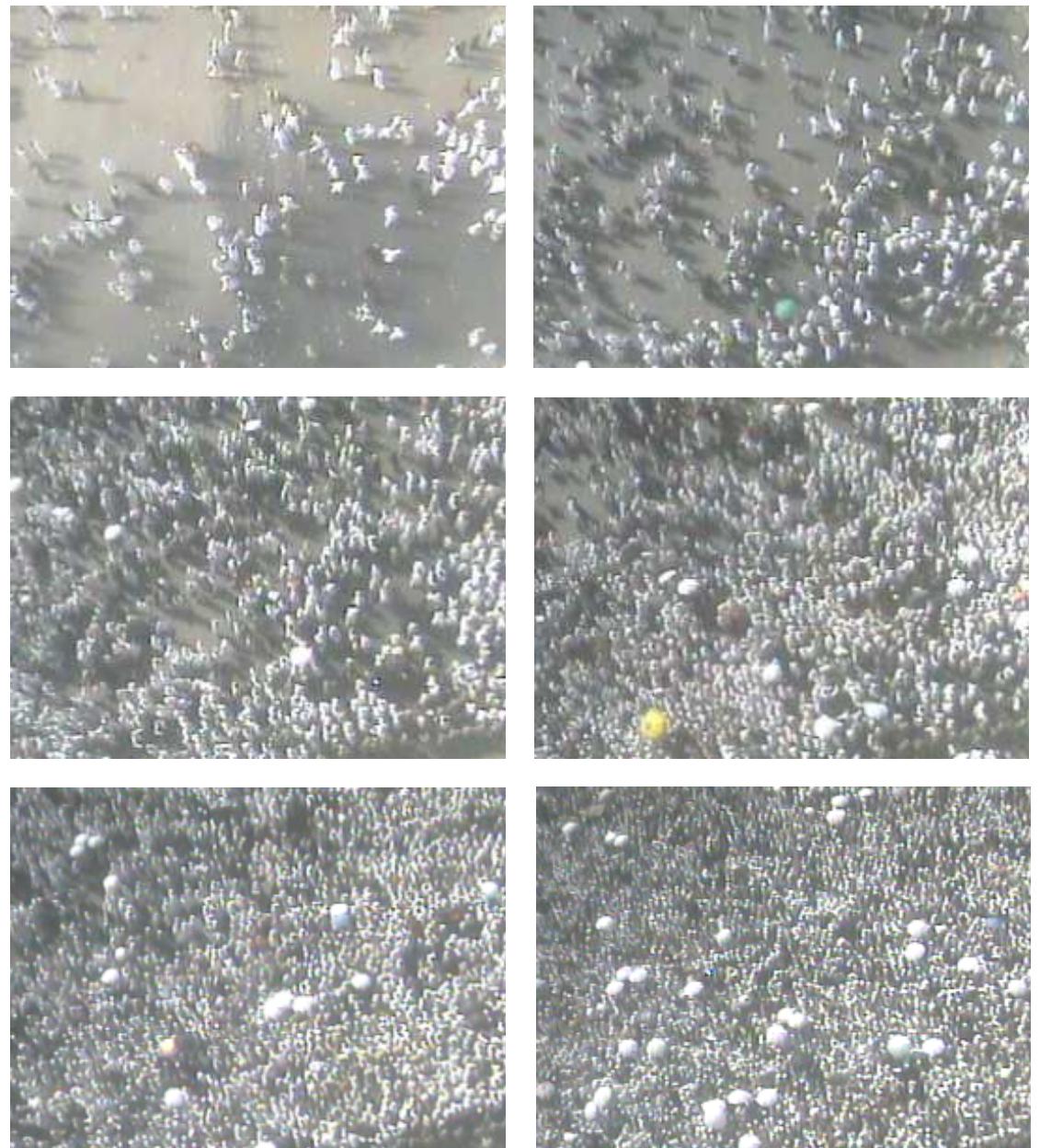


Figure 7.6: Snapshots reflecting the density build-up during 90 minutes before the accident occurred, until the time when crowd turbulence occurred. One can clearly see that the distribution of pedestrians is not homogeneous. The local densities vary considerably (see Sec. 4.2, Fig. 4.1). The bright circular objects are umbrellas, used as protection from sunshine. See Sec. 4.3.1 for how to filter away the effect of these umbrellas from the data.

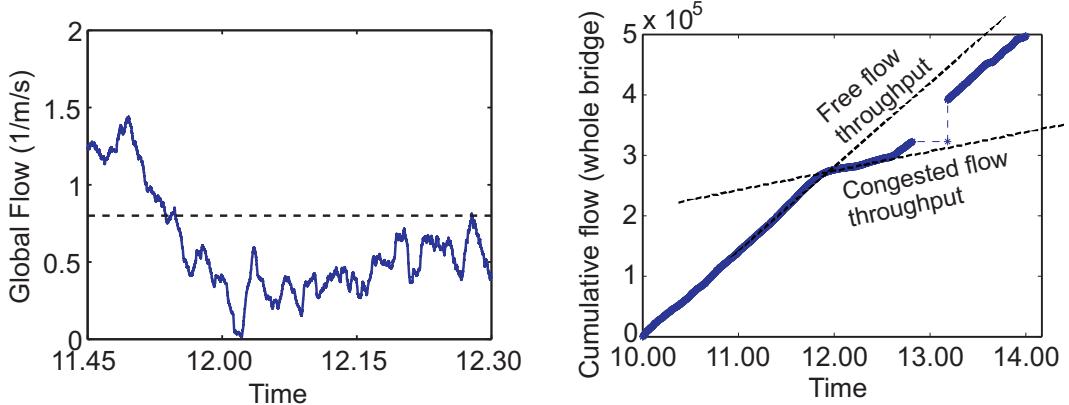


Figure 7.7: Left: During only 15 minutes, the inflow to the Jamarat Bridge turned from a point close to capacity (i.e. the maximum value), all the way down to zero. This transition may have been amplified by religious rituals. However, the flow did not stay at zero, but started to oscillate, giving rise to stop-and-go flow (see Sec. 3.6). Right: The cumulative inflow to the previous Jamarat Bridge shows that there is a decrease in the flow before and during the crowd disaster, but the average flow never vanishes over significant time periods. The gap in the curve is due to missing data.

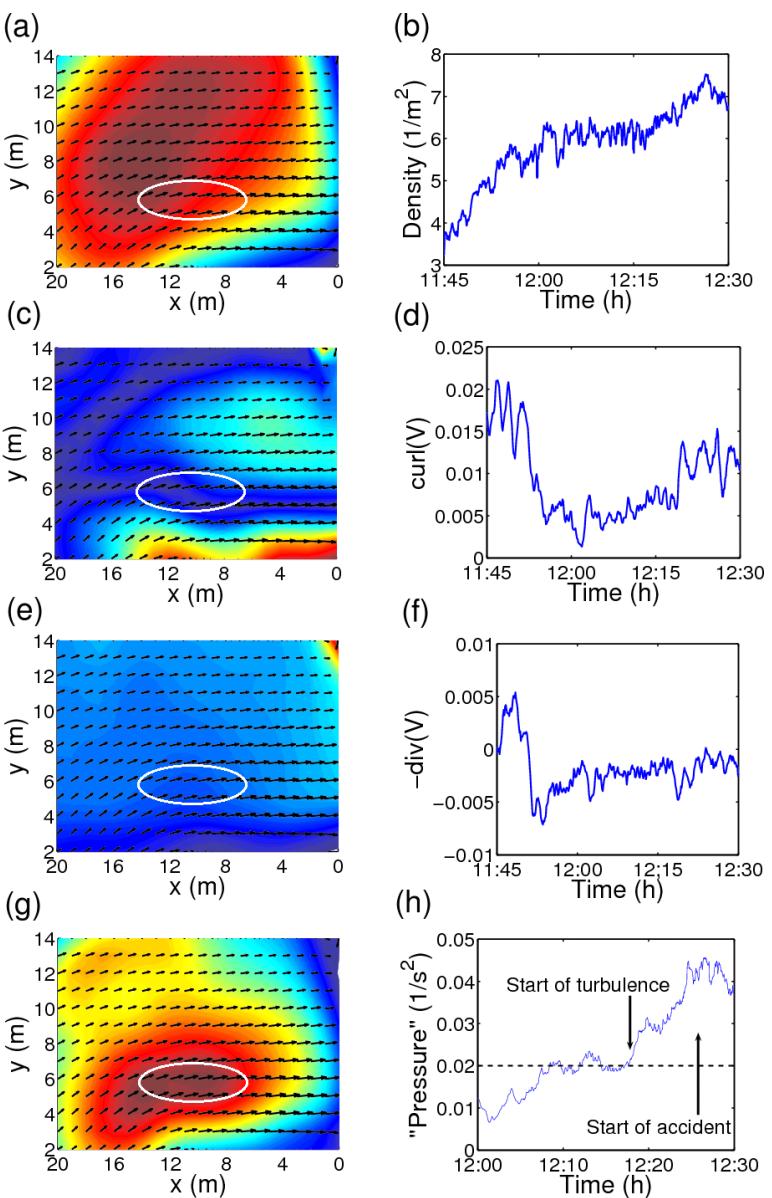


Figure 7.8: Test of various quantities regarding their ability to represent the time periods and areas of critical crowd conditions. The ellipse marks the location where the disaster occurred. The area where the video recording was made, is shown in Fig. 7.2.

- a) Density averaged over time.
- b) Density averaged over space.
- c) Curl $\|curl(\vec{V})\|$, averaged over time.
- d) Curl averaged over space.
- e) Negative velocity divergence $-div(\vec{V})$, averaged over time.
- f) Negative velocity divergence, averaged over space.
- g) Pressure, $P(\vec{r}, t) = \rho(\vec{r}, t)Var_{\vec{r}, t}(\vec{V})$ averaged over time.
- h) Pressure averaged over space. Only the pressure and density were useful to represent critical crowd conditions in time. However, considering the location of the disaster, only the pressure measurement reflected where the disaster occurred.

7.5 Large-Scale Evaluation of Videos

After the crowd disaster, the video evaluation software (see Chapter 2 for details) was developed, to analyze the crowd disaster, and it was then used during the Hajj in 1427H (December 30, 2006 until January 1, 2007) to monitor the flows. This served as useful input for the security's crowd management decisions regarding the times when pilgrim groups should start their tours and the streets they should take.

The system (see Appendix C for details) was set up and run on 24 video cameras mounted primarily on and around the Jamarat Bridge (see Fig. 7.9). The cameras were mounted in two layers, one covering all input ramps and all output ramps to the bridge itself, and another layer covering all the streets and stairways leading to and from the Jamarat Plaza (the area around the Jamarat Bridge).

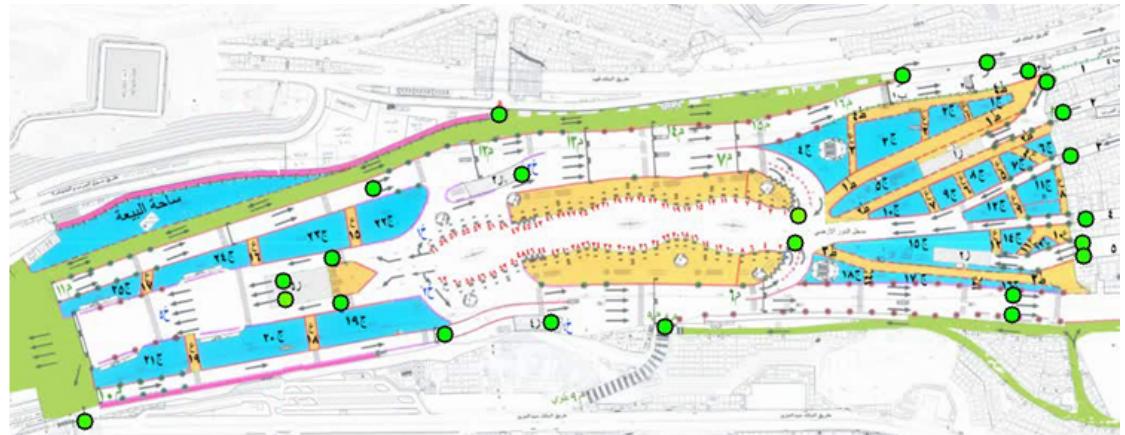


Figure 7.9: Camera positions (green circles) during the Hajj in 1427H. The Jamarat bridge was reconstructed after the Hajj in 1426H, which led to new camera positions as well.

With cameras covering all the inputs and outputs to and from the two areas, we were able to add all the flows of pilgrims and determine the difference between inflows and outflows, and from this information the increase or decrease of the crowd density inside of each area can be calculated. Further on, we could compare the pilgrim flows under each camera to the available capacity of the corresponding section. By this, the security forces were able to assess the potential risk of over-utilization of different parts of the system [77]. This information supported their decisions to stop or redirect flows of pilgrims depending on their spatial distribution.

See Figs. 7.10 and 7.11 for details of how the results were shown.

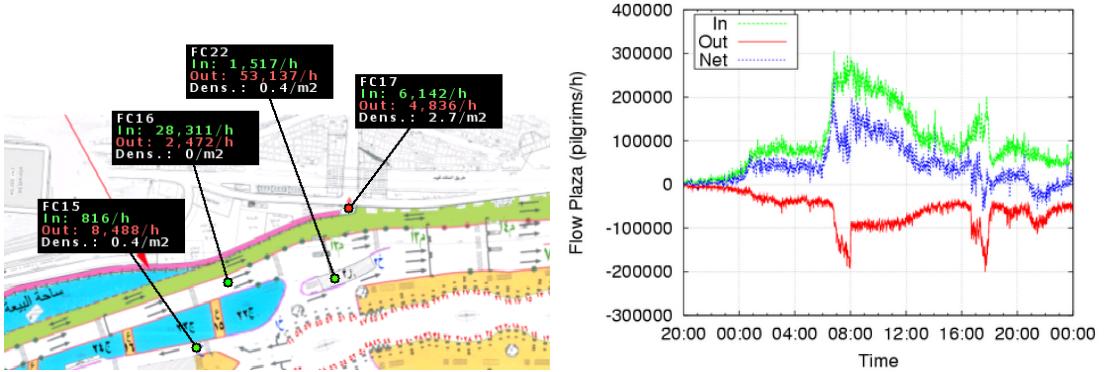


Figure 7.10: Graphical interfaces used during the Hajj in 1427H. Left: Every camera is shown on a map of the area where it was mounted, and updated values of crowd density and flow in both directions were continuously shown. Right: Time-series of the flow and cumulative flow were continuously updated and shown, both for the Jamarat Plaza and for the Jamarat Bridge.

In order to assess the accuracy of the video-tracking software, the total number N_a of pilgrims, automatically counted with the computer software was compared to the total number N_m of pilgrims from the manual counting, determined after the Hajj. The error was defined as $(N_a - N_m)/N_m$. Five cameras, which were selected for the accuracy assessment, are shown in Fig. 7.12. The comparison was made for 72 hours of video recordings. The automatic counts were made in real-time, during the pilgrimage, and the manual counts performed afterwards, from the video recordings of the same cameras. The corresponding errors are displayed in Table 7.1.

Naturally, some differences between manual and automated counts were found. Part of the difference comes from the method of the manual counts, as only 1.2% of the pilgrims were counted (counting was performed during 5 seconds, every 10 minutes). However, in order to see if there is any remaining error, i.e. a bias towards either overestimating or underestimating the number of pilgrims, we have added the cumulative flows of 10 cameras together (and divided it by 10, to get an average number of pilgrims per camera). When comparing these aggregated cumulative flows with the manual counts (Fig 7.13), we see that there is neither any significant underestimation nor any significant overestimation of the number of pilgrims. Therefore, the remaining error seems to go towards 0, the more cameras are added. This suggests that the error is of statistical,

Camera name	Number of pilgrims, N_m	Error
Video 1	800,000	-6%
Video 2	875,000	-6%
Video 3	1,200,000	-10%
Video 4	1,100,000	-15%
Video 5	1,900,000	+13%

Table 7.1: Total number of pilgrims and corresponding relative differences between manual and automated counts, for five different cameras. Snapshots of each of the cameras can be seen in Fig. 7.12

rather than of systematic nature.

A positive side effect of the system being run during the operation of the Hajj is that there is a large amount of data (24 cameras, 3 days, and 2 samples per second), which can be used to analyze different aspects of the Hajj.

The first thing we do with the data is to accumulate all the cameras that were monitoring streets leading to and from the first floor of the Jamarat Bridge. By doing so, we can get an estimate of the total number of pilgrims who were participating in the Hajj. See Fig. 7.14 for the results. During the Hajj in 1427H, there were two possibilities to perform the stoning, on the ground floor and on the 1st floor. Therefore, the total number of pilgrims who participated in the Hajj in 1427H, was roughly estimated by doubling the value measured for the 1st floor.

When comparing the total inflow to the Jamarat Bridge to the total outflow from the bridge, one can find a time offset corresponding to the time a pilgrim spends on the Jamarat Bridge. To determine it, the cross-correlation function

$$R(k) = \frac{1}{n-k} \sum_{t=1}^{n-k} (X_t^{in} - \mu)(X_{t+k}^{out} - \mu) \quad (7.1)$$

was used, where X^{in} is the time-series of the inflow to the Jamarat Bridge and X^{out} is the time-series of the outflow, μ is the mean flow, n the number of samples, and k the time offset. The same procedure was repeated also for the inflows at the Eastern boundary to the Jamarat Plaza. The results are shown in Fig. 7.15. It turns out that pilgrims typically spend 18 minutes on the Jamarat Bridge, and they spend 42 minutes on the Jamarat Plaza. Note, however, that there is a large distribution of these values, since pilgrims show a wide variation of habits in how to perform their pilgrimage in

terms of stoning the Jamarahs, praying, sight-seeing and so on.

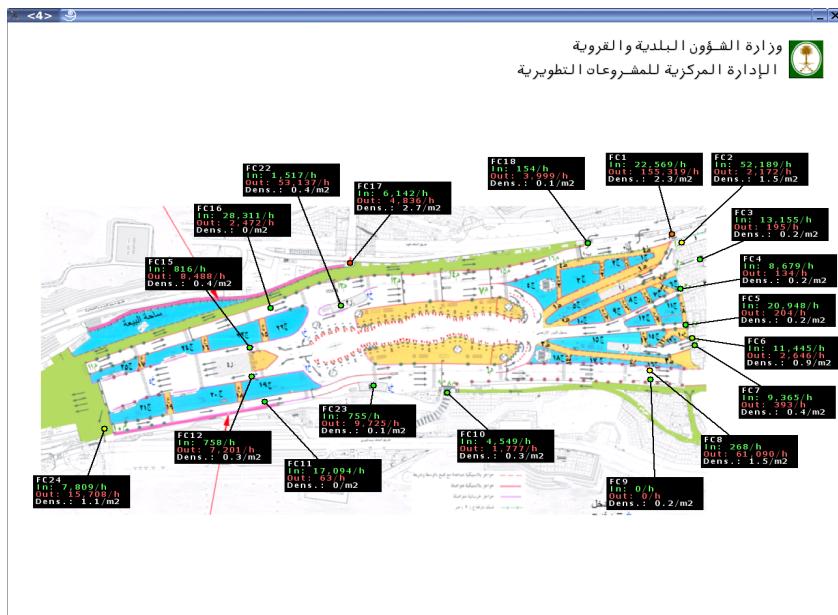
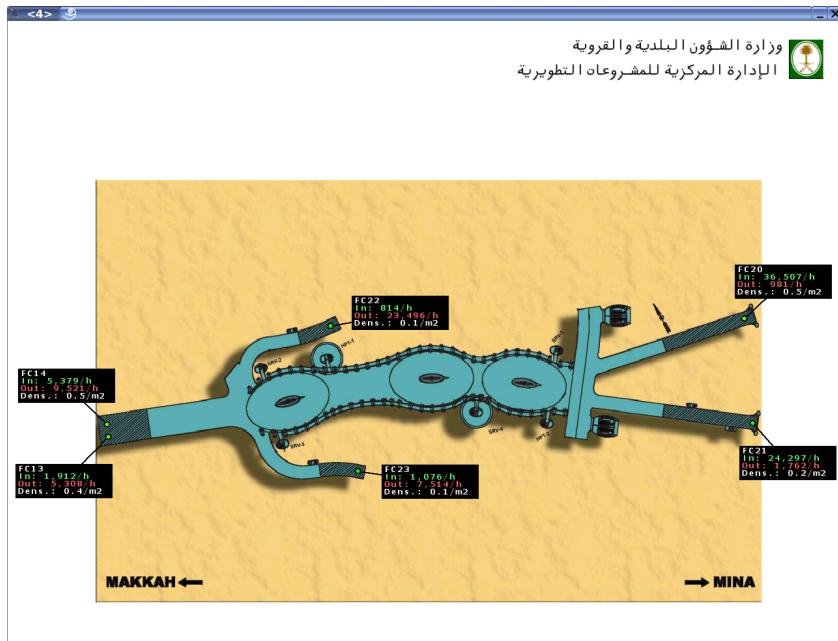
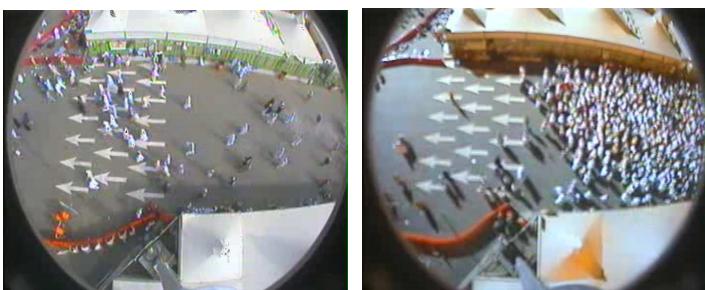


Figure 7.11: Snapshots of the graphical interfaces showing the current state of the system. Top: First level of the new Jamarat Bridge. Bottom: Jamarat Plaza.

Video 1



Video 2



Video 3



Video 4



Video 5



Figure 7.12: Snapshots of five video cameras, mounted onto and around the Jamarat Bridge in Mina, Kingdom of Saudi Arabia, during Hajj 1428H. Each row shows two snapshots of the same camera, with different levels of crowd density. Table 7.1 shows the corresponding measurement errors.

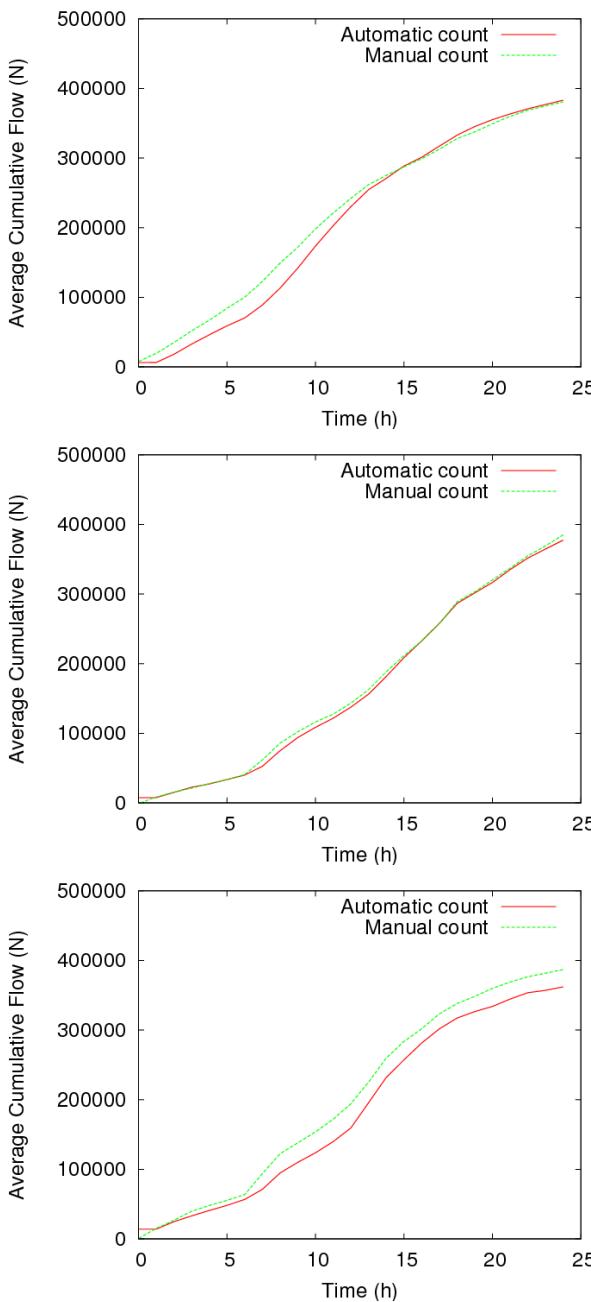


Figure 7.13: Aggregated values when adding the cumulative flows, measured for 10 cameras together, and divided by 10 to get average cumulative flows. Top: 10th of Dhu al-Hijjah with a relative error of 0.2%. Middle: 11th of Dhu al-Hijjah, with a relative error of -2%. Bottom: 12th of Dhu al-Hijjah, with a relative error of -6%

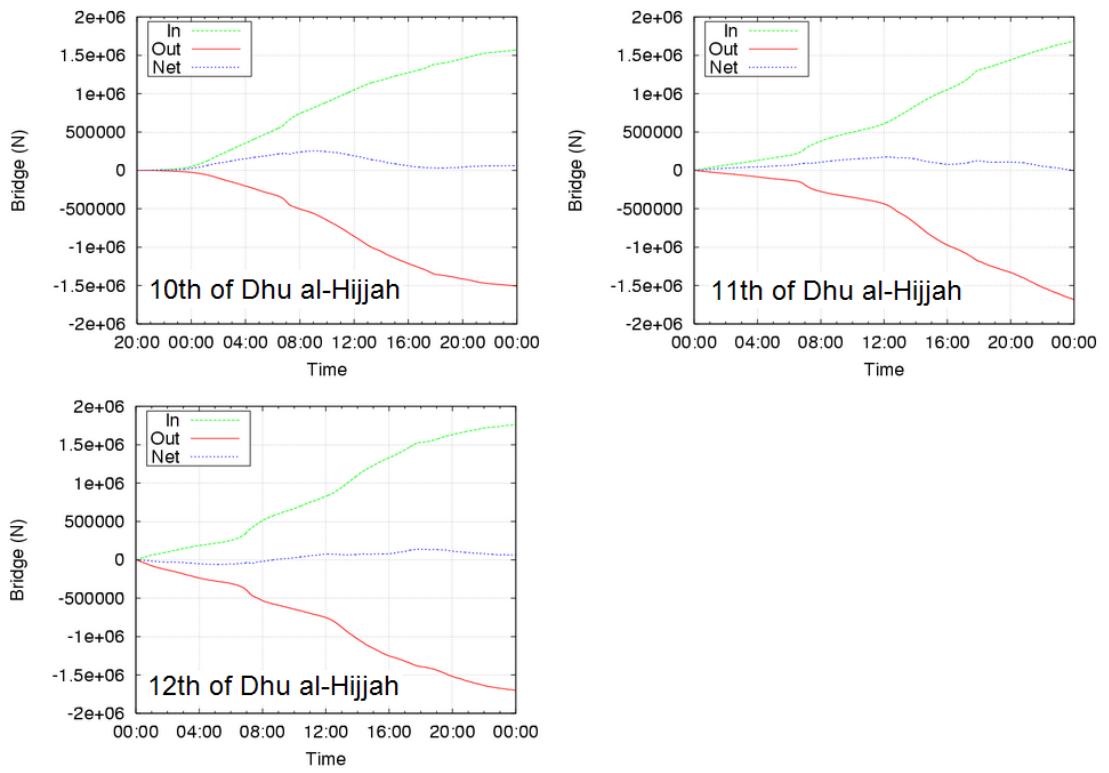


Figure 7.14: Cumulative flow of pilgrims to and from the 1st floor of the Jamarat Bridge, during the Hajj in 1427H.

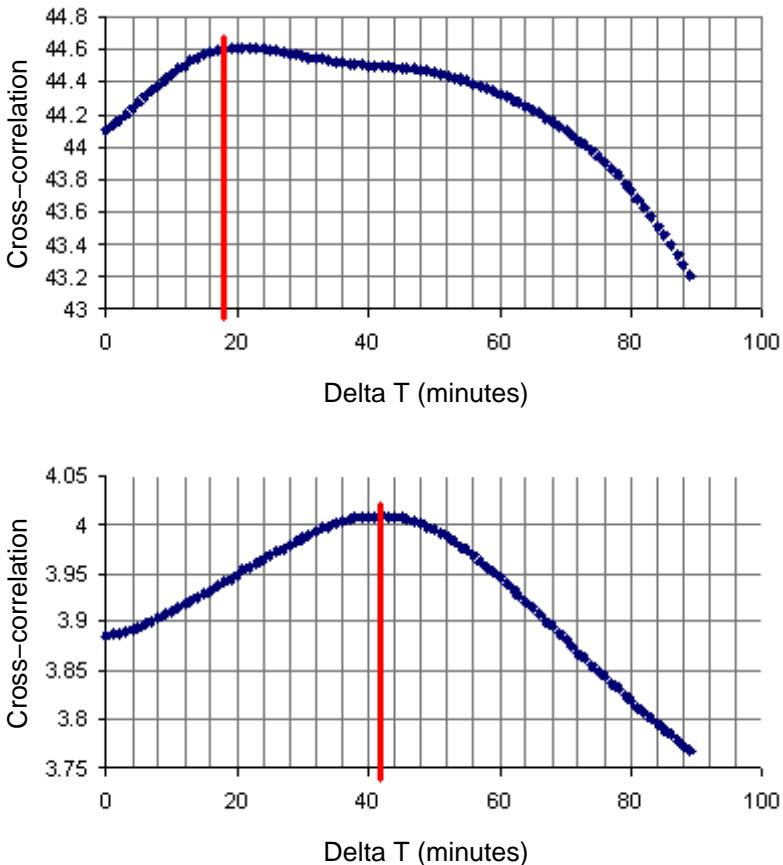


Figure 7.15: Cross-correlation of the inflow and outflow of pilgrims at the first floor of the new Jamarat Bridge (top) and the Jamarat Plaza (bottom). The maximum values (marked with a red line) correspond to the typical travel times. Therefore, during the Hajj in 1427H, a pilgrim typically spent 18 minutes on the Jamarat Bridge and 42 minutes on the Jamarat Plaza (per day).

8 Conclusions and Outlook

It has been demonstrated how different parts of pedestrian research, such as modeling, video tracking, optimization and applications to real-world scenarios, can be brought together. Part of the previous research has been re-evaluated by means of empirical data and tools that were not available a few years ago. By measuring only things relevant to crowd dynamics and by making use of the rich amount of known properties of pedestrians and crowds, a method was introduced which can simultaneously evaluate thousands of pedestrians from videos.

For some aspects such as the distance-dependence of the repulsive forces between pedestrians, empirical data supported an exponential decay. The angular-dependent interactions turned out to be asymmetric and stronger in front of pedestrians than behind their backs. However, it is surprising that this happens to the degree that pedestrians almost do not react *at all* to what happens behind their backs.

By evaluating various specifications of the social-force model, it turned out that velocity-dependent forces greatly increase the performance of the model, in terms of a better fit to empirical trajectory data. The reason for this is that the velocity dependence reflects the pedestrians anticipatory behavior.

Previously, it was assumed that pedestrians would interact with *all* surrounding pedestrians. We have seen that this is not the case. Rather pedestrians typically interact with their eight closest neighbors only. Also, when it comes to evacuations, it has been shown that when evacuees are socially influenced by their eight closest people, evacuation curves similar to empirical data can be obtained, without having to assume any arbitrary distributions of pre-movement time.

For high-density data, in contrast, new results are significantly different from what has been assumed before. Rather than a smooth pedestrian flow in accordance with the previously measured fundamental diagrams, it turns out that the pedestrian flow enters two consecutive dynamic regimes at high densities.

The stop-and-go waves and turbulent regimes have been reported before at an anecdotal level of detail. Given a unique dataset, including extremely high crowd densities, it

was for the first time possible to make a quantitative study of the observed stop-and-go waves and the subsequent phase of crowd turbulence.

From the analysis of the crowd turbulence and the crowd disaster that followed, a new quantity (the “crowd pressure”) has been identified, which may help to anticipate crowd disasters both in space and time.

Note however, that crowd management needs to react *before* crowd turbulence occurs. When the flow drops or stop-and-go waves appear, inflow must be reduced significantly. In fact, the goal of crowd management is to avoid flow breakdown which requires to operate the system considerably *below* capacity, taking into account safety margins.

In the field of traffic engineering, it is usual practise to come up with fit functions as regressions of empirical data.

By analyzing fundamental diagrams from different authors, it was found that all of them shared the property that there was a constant limit of the net-time headway for increasing densities. This property could then be used to show how the shape of the fundamental diagram follows naturally, without having to come up with any arbitrary fit functions, as has been made in the past.

Even though pedestrian research is now actively carried out for many decades, there are still many interesting directions to explore.

Currently, much work is going on in fields such as video-tracking of pedestrians and the analysis of bottleneck flows as well as agent-based modeling.

Many parts of pedestrian modeling are still in an early stage, for example the consideration of decision making and learning, the establishment of norms, and multi-modal transportation, where pedestrians are part of a larger complex system.

Furthermore, in the future, there should be an interdisciplinary collaboration between different research communities. For example:

- Instead of using Kalman or Particle Filters for the video-tracking of pedestrians, one could use specific models of pedestrian motion from the modeling community.
- Fusing simulations with video-tracking would enable an adaptive forecasting of pedestrian flows.
- One should go beyond regressions when analyzing empirical data, and make use of the tools from Pattern Recognition and Machine Learning, or fuse them with simulations, as was done in Chapter 5.
- Cooperation should be initiated between the modeling community, psychologists

and social scientists to develop more human-like, rather than particle-like, models of pedestrian behavior.

It has been demonstrated that many of the self-organization phenomena can be reproduced with simple stimulus-response models. Now the challenge is to make models with even stronger *emergent* properties, e.g. agent-based models where the agents can perform some cognitive processing, having memories, being able to communicate and cooperate, and being adaptive to the environment.

In the past, and still today to some extent, models were designed for fitting the data that they were supposed to explain. For reproducing new features or phenomena, additions were made to the model. Much more interesting would be to make general models, letting them run, and see how they evolve. Then we could learn something significantly different from the designed properties of the model.

A Glossary

A glossary of specific terms used in this thesis follows here. Note that many of these terms have a wider meaning, but this glossary is being tailored to the area of pedestrian research:

- **Crowd density:** (or shortly 'Density'): The average number of pedestrians per m^2 within an area of reference. Unit: m^{-2} .
- **Flow rate:** (or shortly 'Flow'): The number of persons passing the cross section of a road segment per second. Unit: s^{-1} .
- **Specific flow rate:** (or shortly 'Specific flow'): A normalized flow, obtained by dividing the flow with the width of the road segment, where it is being measured. Unit: $m^{-1}s^{-1}$) [persons per meter per second].
- **Fundamental diagram:** The relation of the specific flow as a function of crowd density.
- **Capacity:** The crowd density giving the maximum flow, i.e. the density at the peak of the fundamental diagram.
- **Crowd:** A collection of pedestrians. They may or may not belong to the same group and they may or may not have the same goals, e.g. the same desired direction of motion and the same desired walking speed. Sometimes, the words *pedestrian* and *crowd* are being used as synonyms, but the idea is that the word *pedestrian* implies that the current focus is on interactions between pedestrians, where *crowd* on the other hand, is used when focus is moved from the individuals to the group as a whole.
- **Social-force model:** A simulation model of pedestrians that is based on superposition of forces applied to each pedestrian (particle). These forces include: adaptation of the current walking velocity to a desired velocity, repulsive and attractive social forces between pedestrians and social and physical repulsive forces between pedestrians and obstacles.

- **Bottleneck:** A location where the cross section of the walkable space is decreased, giving a decrease in the flow rate, if the crowd density is high enough. A particular case of a bottleneck is a door through which pedestrians have to pass when they are evacuated from a building.
- **Modeling:** A simplified representation of the reality, made in a theoretical and formal way.
- **Simulation:** An implementation of a model which is being run on a computer. In case of pedestrian research, the simulations are typically investigating the fitness of architectural design patterns or evacuation from buildings, both during normal conditions and conditions of panic and/or fire.
- **Calibration:** The procedure of minimizing the difference of the results obtained by simulations compared to empirical/experimental data. The fitting procedure is normally performed by adjusting the parameters of the model, giving the smallest possible deviation. However, if the error remains too large, it may be needed to change the functional relationships or adding new components to the model as well.
- **Video Tracking:** Automatic tracking of objects (pedestrians) from videos.

B Parameters and Abbreviations

Abbreviation	Description
ANN	Artificial Neural Network
PCA	Principal Components Analysis
SVM	Support Vector Machine
DVR	Digital Video Recorder
PDF	Probability-Density Function
FFT	Fast Fourier Transform
DCT	Discrete Cosine Transform
FPS	Frames Per Second
(J)PDAF	(Joint) Probability Data Association Filter

Table B.1: A list of abbreviations.

Variable	Unit	Description
\vec{r}	m	Pedestrian position
\vec{v}	m/s	Pedestrian velocity
\vec{a}	m/s^2	Pedestrian acceleration
ϱ	m^{-2}	Global crowd density
ρ	m^{-2}	Local crowd density
Q_s	$\text{m}^{-1}\text{s}^{-1}$	Specific Flow
Q	s^{-1}	Flow
\hat{T}	s	Net-time headway

Table B.2: A list of measurables.

Parameter	Unit	Description
<i>Fundamental diagram parameters</i>		
R	m	Radius for measuring local quantities
γ		Fraction of densities to remove for filtering out umbrellas
p		Fraction of umbrellas
<i>Video-analysis parameters</i>		
r'		Average head radius
T'		Threshold value for the Sobel transform
γ'		Fraction-of-head parameter, used together with the Hough transform
E		Relative flow error, between automatic and manual measurements
Δt	s	Time step
<i>Social-force model parameters</i>		
A	m/s^2	Interaction strength
B	m	Distance scaling
N		Number of interactions
τ	s	Relaxation time
λ		Anisotropy constant
ΔT	s	Anticipation time
SP		Speed percentile
T^d	s	Delay
T	s	Forecasting time for calibrating the social-force model
<i>Flow optimization by Boolean grids</i>		
c		Number of times to sequentially apply the clustering rule
d		Neighborhood size used by the clustering rule
k		A constant, defining the level of stochasticity of the clustering rule

Table B.3: A list of parameters.

C Technical Details for Video Analysis

For the video analysis software introduced in Chapter 2 and applied on the Hajj (see Chapter 7), there are various technical details which will follow here.

In order to assess the accuracy of determining crowd density, the following process is made:

1. AVI files are created from the Digital Video Recorders (DVR) used for recording the videos during the Hajj. It would take 24 hours of processing to create 24 hour of AVI videos.
2. From each 24 hours of video recordings, 30 snapshots are created at different times and different levels of crowd density.
3. In each of these snapshots, a rectangle is drawn, with arbitrary size and location.
4. Each of these 30 snapshots are inspected by three persons, and each person manually counts the number of heads inside the rectangle (see Fig. 7.1).
5. A calibration process is started, to automatically find the optimal parameter values matching the 30 snapshots (15 of the snapshots are used as a training dataset and the other 15 snapshots are used as a validation dataset).
6. The optimal parameters are given to the video-tracking software, and the software measures the flows during 24 hours of video.
7. The same video is played on a large-screen TV, and the pedestrian flows are counted manually (see Fig. C.3).
8. A validation (comparison) is made between the automatic measurements and the manual counts. Figure 7.3 shows the result for one video during 24 hours.

C.1 Technical Setup

For each camera, the following procedure was undertaken:

- Mechanical calibration: A 30 m crane was used to set the camera angle and zoom to fit the desired area well. The communication was made over a mobile phone, with one person in front of the computer, one person in the crane and two persons on the ground under the camera. The two persons on the ground were used as mobile markers to find the boundary of the viewable area.
- Distance calibration: When the boundary of the camera was fixed, one person made two dots with a red spray on the ground, one at each of two edges of the video. Then the distance between the two dots was measured.
- Basic software calibration: The operator defined exactly, where on the video screen the analysis should be performed, set the parameters for the lens, and defined the distance in meters between the two red dots on the screen.
- Advanced software calibration: To get reliable values of the speeds, flows and densities, the operator had to mark some heads moving under the camera. This procedure had to be performed two times, once at daytime and once at night. See Fig. C.1 for details.

Since there are no pilgrims in the area around the Jamarat Bridge during the 8th and 9th of Dhu al-Hijjah (the two days before the pilgrims go to the Jamarat Bridge), a group of soldiers was used for calibration, during the Hajj in 1427H. Specifically, 200 soldiers (see Fig. C.2) were moving under each of the cameras, in order to gain data for the calibration of the software.

The measurements made by the video-analysis software include:

- Average velocity for two directions (typically along a street in both directions).
- Average flow in both directions.
- Cumulative flow in both directions.
- Average crowd density.

An SQL database was run on each computer, in which all data entries described above were stored. The resolution was around 2 samples per second for each camera. This setup has several good features:

- Speed: SQL databases provides a fast solution both for storing and reading data.

- Flexibility: Once the data has been stored in the database, it can be simultaneously read by a separate application, for example to generate statistics or to display the numbers in a graphical way.
- Communication: The SQL software automatically takes care of the network communication between the different computers.
- Decentralization: If a network link would go down, no significant amount of data would be lost, since the data are primarily stored in the SQL database on the local computer.
- Easiness to aggregate results: A script was produced to continuously read data from all databases, creating 1-minute averages and then writing them to a separate (global) database. Even if this database would be lost for any reason, it would be possible to restart this process and rebuild the global database from scratch during operation of the system.

During operation, the results were displayed according to Fig. C.4.

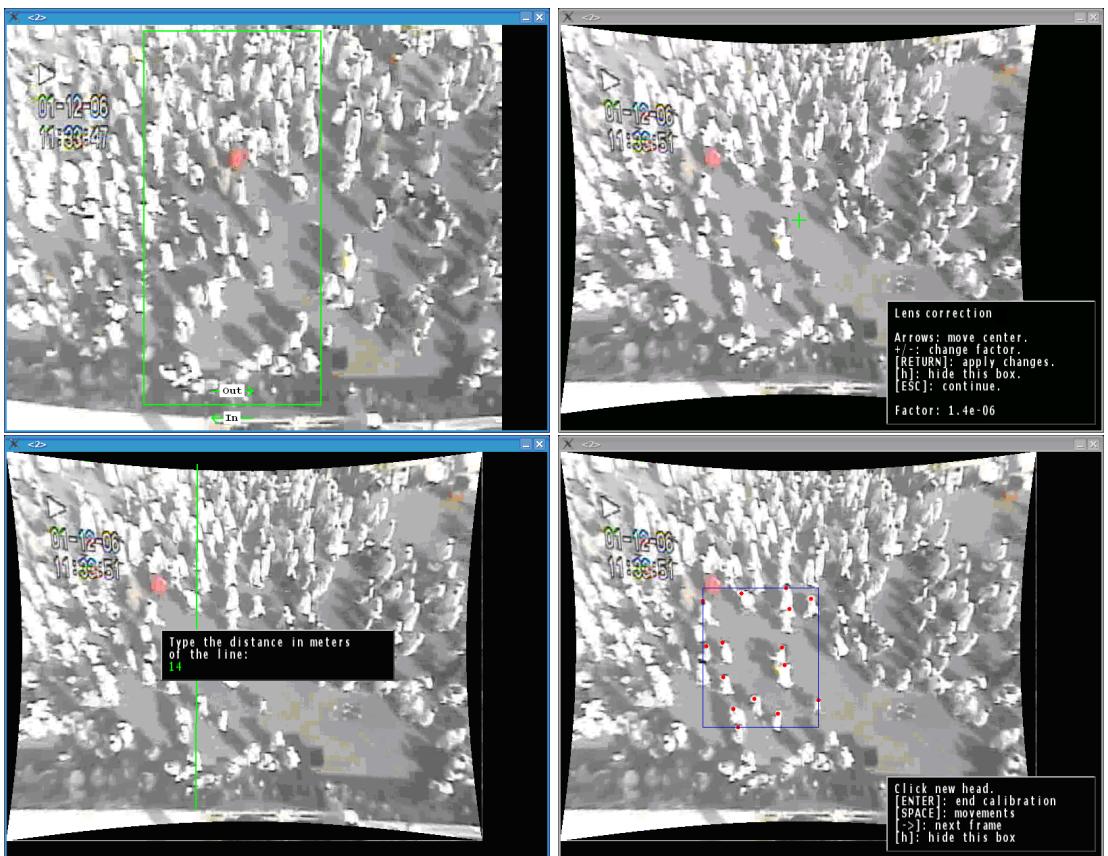


Figure C.1: Snapshots of the calibration screens. Top left: The area of analysis is marked in the video. Top right: The lens is corrected for. Bottom left: The distance calibration is made. Bottom right: Heads of pedestrians are manually clicked with the computer mouse. These data are then used for calibrating the software.



Figure C.2: In order to gain data for performing the calibration, 200 soldiers were moving under each of the cameras.



Figure C.3: Manual flow-measurement process. Note that the screen is divided into four parts to make it easier to keep an overview of each sub-area.

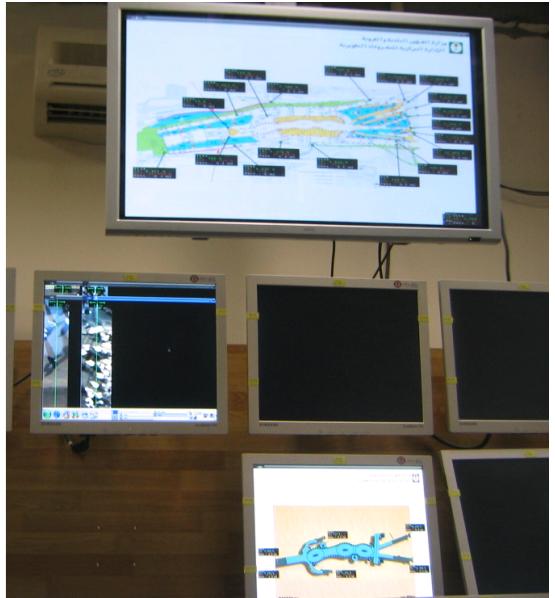
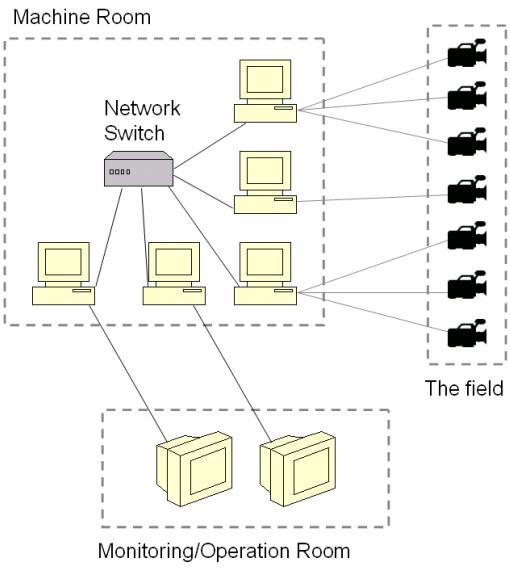


Figure C.4: Left: Cameras are mounted on various locations in the field, and then all the cameras are connected to computers inside a machine room. These computers analyze the videos, aggregate the data and then present the current state of the systems graphically on monitors in the monitoring room. Right: A photo from the monitoring room, showing the graphical interfaces.

D Publications

Peer-Reviewed Journal Papers

1. Dyer, J. R. G., Johansson, A., Helbing, D., Couzin, I. D., and Krause, J. (2009) Leadership, consensus decision making and collective behaviour in human crowds, *Philosophical Transactions of the Royal Society B: Biological Sciences* **364**, 781–789.
2. Moussaid, M., Helbing, D., Garnier, S., Johansson, A., Combe, M., and Theraulaz, G. (2009) Experimental study of the behavioural mechanisms underlying self-organization in human crowds, *Proceedings of the Royal Society B: Biological Sciences*, in print.
3. Helbing, D. and Johansson, A. (2009) On the controversy around Daganzo's requiem for and Aw-Rascle's resurrection of second-order traffic flow models, *European Physical Journal B*, in print.
4. Johansson, A., Helbing, D., Al-Abideen, H. Z., and Al-Bosta, S. (2008) From crowd dynamics to crowd safety: A video-based analysis, *Advances in Complex Systems* **11** (4), 497–527.
5. Nilsson, D. and Johansson, A. (2008) Social influence during the initial phase of a fire evacuation—Analysis of evacuation experiments in a cinema theatre, *Fire Safety Journal* **44** (1), 71–79.
6. Helbing, D., Johansson, A., and Al-Abideen, H. Z. (2007) The dynamics of crowd disasters: An empirical study, *Physical Review E* **75**, 046109.
7. Johansson, A., Helbing, D., and Shukla, P. K. (2007) Specification of the social force pedestrian model by evolutionary adjustment to video tracking data, *Advances in Complex Systems* **10**, 271–288.

8. Yu, W. and Johansson, A. (2007) Modeling crowd turbulence by many-particle simulations, *Physical Review E* **76**, 046105.
9. Helbing, D., Johansson, A., Mathiesen, J., Jensen, M. H., and Hansen, A. (2006) Analytical approach to continuous and intermittent bottleneck flows, *Physical Review Letters* **97**, 168001.
10. Peters, K., Johansson, A., Dussutour, A., and Helbing, D. (2006) Analytical and numerical investigation of ant behaviour under crowded conditions, *Advances in Complex Systems* **9**, 337–352.
11. Helbing, D., Buzna, L., Johansson, A., and Werner, T. (2005) Self-organized pedestrian crowd dynamics: Experiments, simulations, and design solutions, *Transportation Science* **39**(1), 1–24.
12. Peters, K., Johansson, A., and Helbing, D. (2005) Swarm intelligence beyond stigmergy: Traffic optimization in ants, *Künstliche Intelligenz* **4**, 11–16.

Book chapters

13. Helbing, D. and Johansson, A. (2009) Pedestrian, crowd and evacuation dynamics, in: Meyers, Robert A (ed.) *Encyclopedia of Complexity and Systems Science* (Springer, Heidelberg).
14. Helbing, D., Kühnert, C., Lämmer, S., Johansson, A., Gehlsen, B., Ammoser, H., and West G. B. (2009) Power laws in urban supply networks, social systems, and dense pedestrian crowds, Pages: 433–450 in: D. Lane, S. van der Leeuw, D. Pumain, G. West (eds.) *Complexity perspectives in innovation and social change* (Springer, Berlin Heidelberg, New York).
15. Helbing, D., Johansson, A., and Lämmer, S. (2008) Self-organization and optimization of pedestrian and vehicle traffic in urban environments, in: S. Albeverio, et. al. (eds.) *The Dynamics of Complex Urban Systems* (Physica-Verlag).
16. Johansson, A. and Helbing, D. (2006) Crowd dynamics, Pages 449–471 in: B. K. Chakrabarti, A. Chakraborti, and A. Chatterjee (eds.) *Econophysics and Sociophysics – Trends and Perspectives* (Wiley, Weinheim).

17. Helbing, D., Johansson, A., and Buzna, L. (2006) New design solutions for pedestrian facilities based on recent empirical results and computer simulations, Pages 67–88 in: M. Aron, F. Boillot and J.-P. Lebacque (eds.) *Modélisation Du Traffic - Actes du groupe de travail 2003, Actes No. 104*.

Conference Proceedings Papers

18. Johansson, A. and Helbing, D. (2008) Analysis of empirical trajectory data of pedestrians, *Pedestrian and Evacuation Dynamics 2008*, in print.
19. Johansson, A. and Helbing, D. (2007) Pedestrian flow optimization with a genetic algorithm based on Boolean grids, Pages 237–272 in: N. Waldau, P. Gattermann, H. Knoflacher and M. Schreckenberg (eds.) *Pedestrian and Evacuation Dynamics 2005* (Springer, Berlin).
20. Helbing, D., Johansson, A., and Al-Abideen, H. Z. (2007) Crowd turbulence: the physics of crowd disasters, Pages 967–969 in: *The 5th International Conference on Nonlinear Mechanics (ICNM-V)* (Shanghai).
21. Helbing, D. and Johansson, A. (2007) Quantitative agent-based modeling of human interactions in space and time, Pages 623–637 in F. Amblard (ed.) *Proceedings of the 4th Conference of the European Social Simulation Association (ESSA'07)* (Toulouse).
22. Helbing, D., Johansson, A., Al-Abideen, H. Z., and Al-Bosta, S. (2007) Evaluations of video data of crowds around the Jamarat Bridge and safety implications for mass events, *European Conference on Complex Systems*, Dresden, on CD-ROM.
23. Helbing, D., Johansson, A., and Buzna, L. (2005) Pedestrian dynamics and evacuation: Empirical results and design solutions, in: H. S. Mahmassani (ed.) *Flow, Dynamics and Human Interaction. Proceedings of the 16th International Symposium on Transportation and Traffic Theory* (Elsevier Science, Amsterdam).
24. Helbing, D. and Johansson, A. (2005) Sicherheit in Fußgängermengen bei Massenveranstaltungen, Pages 87–130 in: C.-S. Grommek (ed.) Rothenburger Beiträge, *Polizeiwissenschaftliche Schriftenreihe der Fachhochschule für Polizei Sachsen*, Band 23. (Rothenburg/Oberlausitz) [in German].

25. Helbing, D. and Johansson, A. (2005) Grundsätzliche Hinweise zur Theorie und Konzeption von Evakuierung und Räumung: Sicherheit in Fußgängermengen bei Massenveranstaltungen, Pages 1–42 in: *VdS-Fachtagung. Evakuierung und Räumung von Veranstaltungen* (VdS, Köln) [in German].

Other

26. Johansson, A., Chan, S., Hamdar, S., Al-Bosta, S., Haase, K., and Helbing, D. (2008) Towards automatic analysis and forecasting of pilgrim flow by video analysis and simulation, *Specialized Architectural Engineering Magazine issued by Ministry of Municipal and Rural Affairs, Kingdom of Saudi Arabia*, in print.
27. Al-Abideen, H., Z., Al-Bosta, S., Johansson, A., and Helbing, D. (2006) Scientific evaluation of videos showing hajjis flow into Jamarat area, *Specialized Architectural Engineering Magazine issued by Ministry of Municipal and Rural Affairs, Kingdom of Saudi Arabia* **10**, December 2006.
28. Johansson, A. (2004) Pedestrian simulations with the social force model, Master's thesis, Dresden University of Technology, Dresden, Germany.

Bibliography

- [1] Al-Abideen, H. Z. (2005) The development of multistory Jamarat Bridge and surrounding area. 6th International Congress, Global Construction, University of Dundee, Scotland, 5-7 July, 2005.
- [2] Al-Abideen, H., Z. and Al-Bosta, S. (2006) Human crowd management, movement of pilgrims around Jamarat in Mina in Hajj 1426H, Workshop: Management and Control of Human Crowds, Riyadh, Kingdom of Saudi Arabia, 5-8 February, 2006.
- [3] Al-Abideen, H., Z., Al-Bosta, S., Johansson, A., and Helbing, D. (2007) Scientific evaluation of videos showing hajjis flow into Jamarat area, *Specialized Architectural Engineering Magazine issued by Ministry of Municipal and Rural Affairs, Kingdom of Saudi Arabia*, **10**, December 2006.
- [4] Al-Bosta, S. (2006) Consequences of understanding of Jamarat stoning capacity, *Workshop on the results of the program prepared for the scheduling and management of hajjis groups to Jamarat for 1427H Hajj season*, Jeddah, Kingdom of Saudi Arabia, 26-29 November 2006.
- [5] Al-Bosta, S. (2006) Manual for fences for pedestrian paths, *6th Workshop: Design and Organization of Jamarat Plaza during Next Hajj 1427H*, Jeddah, Kingdom of Saudi Arabia, 5-8 February 2006.
- [6] Al-Bosta, S. and Serwill, D. (2006) Design and organization of Jamarat Plaza in Hajj 1427H, *Specialized Architectural Engineering Magazine issued by Ministry of Municipal and Rural Affairs, Kingdom of Saudi Arabia* **10**, December 2006.
- [7] Al-Bosta, S. (2007) Evaluation on the squatter and luggage situation in Mina during Hajj 1427H, *Workshop dealing with squatters and luggage in Hajj 1428H*, Makkah, Kingdom of Saudi Arabia, 23-25 April 2007.
- [8] AlGadhi, S.A.H., Mahmassani, H.S., and Herman, R. (2002) A speed-concentration relation for bi-directional crowd movements with strong interaction.

Pages 3–20 in: M. Schreckenberg and S. D. Sharma (eds.) *Pedestrian and Evacuation Dynamics* (Springer, Berlin).

- [9] Ando, K., Ota, H. and Oki, T. (1988) Forecasting the flow of people, *Railway Research Review* **45**, 8-14. (in Japanese).
- [10] Antonini, G. (2005) A Discrete Choice Modeling Framework for Pedestrian Walking Behavior with Application to Human Tracking in Video Sequences, PhD thesis, Ecole Polytechnique Fédérale de Lausanne Switzerland.
- [11] Antonini, G., Bierlair M. and Weber M. (2006) Discrete choice models of pedestrian walking behavior, *Transportation Research Part B*, **40** (8), 667–687.
- [12] Bartels, R., Beatty, J., and Barsky, B. (1987) An Introduction to Splines for use in Computer Graphics and Geometric Modeling (Morgan Kaufmann).
- [13] Bar-Shalom, Y. and Fortmann, T. (1988) Tracking and Data Association (Academic Press).
- [14] Bayer, K., Rejnö, T. (1999) Evacuation alarm, Optimizing through full-scale experiments (in Swedish), Report 5053, Department of Fire Safety Engineering, Lund University, Lund.
- [15] Bazzani, A., Giorgini, B., Servizi, G. and Turchetti, G. (2003) A chronotopic model of mobility in urban spaces, *Physica A: Statistical Mechanics and its Applications* **325**(3-4), 517–530.
- [16] Bertozzi, M., Broggi, A., Grisleri, P., Tibaldi, A. and Del Rose, M. (2004) A tool for vision based pedestrian detection performance evaluation, *Intelligent Vehicles Symposium*.
- [17] Bishop, C. M (2006) Pattern Recognition and Machine Learning. (Springer, Singapore).
- [18] Black, M. and Anandan, P. (1993) A framework for the robust estimation of optical flow, Pages 231–236 in: *In Proc. 4th Int. Conf. on Computer Vision*, Berlin, Germany.
- [19] Blake, A. and Isard, M. (1994) 3D position, attitude and shape input using video tracking of hands and lips, *In Proc. Siggraph*, 185–192.

- [20] Blue, V. J., Adler, J. L. (2000) Cellular automata microsimulation of bi-directional pedestrian flows, *Transportation Research Record* **1676** 135.
- [21] Bolay, K. (1998) Nichtlineare Phänomene in einem fluid-dynamischen Verkehrsmodell, Diploma thesis, University of Stuttgart.
- [22] Borgers, A. and Timmermans, H. J. P. (1985) City centre entry points, store location patterns and pedestrian route choice behavior: a microlevel simulation model, *Socio-Economic Planning Science* **20**, 25–31.
- [23] Bregler, C. and Malik, J. (1998) Tracking people with twists and exponential maps, *In: Proc. CVPR*.
- [24] Brown, R. (1965) *Social Psychology* (Free Press, New York).
- [25] Chowdhury, D., Guttal, V., Nishinari, K., Schadschneider, A. (2002) A cellular-automata model of flow in ant trails: non-monotonic variation of speed with density *J. Phys. A: Math. Gen.* **35** (41), L573–L577.
- [26] Daamen, W. (2004) Modelling Passenger Flows in Public Transport Facilities, PhD thesis, TU Delft.
- [27] Deb, K., Goel, T. (2001) A hybrid multi-objective evolutionary approach to engineering shape design, in: *Proceedings of the First International Conference on Evolutionary Multi-Criterion Optimization (EMO-2001)*, Dortmund, Germany, 385.
- [28] De Marchi, S. (2005) *Computational and Mathematical Modeling in the Social Sciences* (Cambridge University Press, New York).
- [29] Doering, C. R., Gibbon, J. D. (1995) Applied Analysis of the Navier-Stokes Equations (Cambridge University Press, Cambridge).
- [30] Donders, F. C. (1868) On the speed of mental processes. Translated by W. G. Koster, 1969. *Acta Psychologica*, **30**, pp. 412–431.
- [31] Bonabeau, E., Dorigo, M., and Theraulaz, G. (1999) Swarm Intelligence: From Natural to Artificial Systems (Oxford University Press, Oxford).
- [32] Dussutour, A., Fourcassié V., Helbing, D., and Deneubourg, J.-L. (2004) Optimal traffic organization in ants under crowded conditions, *Nature* **428**, 70.

- [33] Eissfeldt, E. and Wagner, P. (2003) Effects of anticipatory driving in a traffic flow model, *The European Physical Journal B - Condensed Matter* **33**(1), 121–129.
- [34] Escobar, R. and de la Rosa, A. (2003) Architectural design for the survival optimization of panicking fleeing victims, in: *7th European Conference on Artificial Life (ECAL 2003)*, Dortmund, Germany.
- [35] Farkas, I., Helbing D., and Vicsek, T. (2002) Mexican waves in an excitable medium, *Nature* **419**, 131–132.
- [36] Fruin, J. J. (1971) Designing for pedestrians: A level-of-service concept, *Highway Research Record* **355**, 1–15.
- [37] Gavrila, D. M. and Giebel, J. (2002) Shape-based pedestrian detection and tracking, In: *Proc. of the IEEE Intelligent Vehicles Symposium*, Paris, France.
- [38] Ghashghaie, S., Breymann, W., Peinke, J., Talkner. P. and Dodge, Y. (1996) Turbulent cascades in foreign exchange markets, *Nature* **381**, 767.
- [39] Gutenberg, B. and Richter, C. F, (1944) Frequency of earthquakes in California, *Bulletin of the Seismological Society of America* **34** 185–188.
- [40] Green, M. (2000) How Long Does It Take to Stop? Methodological analysis of driver perception-brake times, *Transport. Hum. Factors* **2** pp. 195–216.
- [41] Gribbon, K. T. Johnston, C. T. and Bailey, D. G. (2003) A real-time FPGA implementation of a barrel distortion correction algorithm with bilinear interpolation, *Image and Vision Computing NZ*, Palmerston North.
- [42] Halabi, W. S. (2006) Overcrowding and the Holy Mosque, PhD thesis, University of Newcastle upon Tyne.
- [43] Hampapur, A., Brown, L., Connell, J., Ekin, A., Haas, N., Lu, M., Merkl, H., Pankanti, S., Senior, A., Shu, C.-F. and Tian, Y.L (2005) Smart Video Surveillance: exploring the concept of multiscale spatiotemporal tracking, *IEEE Signal Processing Magazine*, 38–51.
- [44] Hanking, B. D. and Wright, R. A. (1958) Passenger flow in subways, *Operational Research Quarterly* **9**, 81–88.

- [45] Harris, C. (1992) Tracking with rigid models, In: Blake, A. and Yuille, A., (eds.), *Active Vision*, 59–74. MIT.
- [46] Helbing, D. (1991) A mathematical model for the behavior of pedestrians, *Behavioral Science* **36**, 298–310.
- [47] Helbing, D. (1992) A fluid-dynamic model for the movement of pedestrians, *Complex Systems* **6**, 391–415.
- [48] Helbing, D. (1994) A mathematical model for the behavior of individuals in a social field, *Journal of Mathematical Sociology* **19** (3), 189-219.
- [49] Helbing, D. (1997) *Verkehrs dynamik. Neue physikalische Modellierungskonzepte* (Springer, Berlin).
- [50] Helbing, D. (2001) Traffic and related self-driven many-particle systems, *Reviews of Modern Physics* **73**, 1067–1141.
- [51] Helbing, D., Buzna, L., Johansson, A., and Werner, T. (2005) Self-organized pedestrian crowd dynamics: Experiments, simulations, and design solutions, *Transportation Science* **39**, 1–24.
- [52] Helbing, D., Farkas, I. J., Molnár, P., and Vicsek, T. (2002) Simulation of pedestrian crowds in normal and evacuation situations. Pages 21-58 in: M. Schreckenberg and S. D. Sharma (eds.) *Pedestrian and Evacuation Dynamics* (Springer, Berlin).
- [53] Helbing, D., Farkas, I., and Vicsek, T. (2000) Simulating dynamical features of escape panic, *Nature* **407**, 487–490.
- [54] Helbing, D., Johansson, A. and Al-Abideen, H. Z. (2007) The Dynamics of Crowd Disasters: An Empirical Study, *Physical Review E* **75**, 046109.
- [55] Helbing, D., Johansson, A., Mathiesen, J., Jensen, M. H. and Hansen, A. (2006) Analytical approach to continuous and intermittent bottleneck flows, *Physical Review Letters* **97**, 168001.
- [56] Helbing, D., Keltsch, J., and Molnár, P. (1997) Modelling the evolution of human trail systems, *Nature* **388**, 47–50.

- [57] Helbing, D. and Molnár, P. (1995) Social force model for pedestrian dynamics, *Physical Review E* **51**, 4282–4286.
- [58] Helbing, D., Molnár, P., Schweitzer, F. (1994) in: *Evolution of Natural Structures (Sonderforschungsbereich 230)*, Stuttgart, Germany, 229–234.
- [59] Helbing, D., Molnár, P., Farkas, I., Bolay, K. (2001) Self-organizing pedestrian movement, *Environment and Planning B* **28**, 361–383.
- [60] Helbing, D., Schweitzer, F., Keltsch, J., Molnár, P. (1997) Active walker model for the formation of human and animal trail systems, *Physical Review E* **56**, 2527.
- [61] Henderson, L. F. (1971) The Statistics of Crowd Fluids, *Nature* **229**, 381–383.
- [62] Highway Capacity Manual, Chapter 13. (Transportation Research Board, Special Report 209, Washington, D.C., 1985).
- [63] Hirasaki, E., Moore, S. T. Raphan, T. and Cohen, B. (1999) Effects of walking velocity on vertical head and body movements during locomotion, *Exp Brain Res* **127**, 117–130.
- [64] Hoogendoorn, S.P. (2002) Walker Behaviour Modelling by Differential Games, in *Proceedings of the Computational Physics of Transport and Interface Dynamics Seminar*, (Springer, 2002).
- [65] Hoogendoorn, S. P. and. Bovy, P. H. L (2004) Pedestrian route-choice and activity scheduling theory and models, *Transportation Research Part B* **38** (2), 169–190.
- [66] Hoogendoorn, S. P. and Daamen, W. (2005) Pedestrian behavior at bottlenecks, *Transportation Science* **39**(2), 147–159.
- [67] Hoogendoorn, S. P., Daamen, W. and Landman, R. (2007) Microscopic calibration and validation of pedestrian models – Cross-comparison of models using experimental data, Pages 253–265 in: N. Waldau, P. Gattermann, H. Knoflacher and M. Schreckenberg (eds.) *Pedestrian and Evacuation Dynamics 2005* (Springer, Berlin).
- [68] Hoogendoorn, S. P., Daamen, W., Bovy, P. H. L. (2003) in: *Pedestrian and Evacuation Dynamics, E. R. Galea (ed.)*, 89–100 (CMS Press, London).
- [69] Hoogendoorn, S. P., Daamen, W., Bovy, P. H. L. (2003) Extracting microscopic pedestrian characteristics from video data, *Annual Meeting Transportation Res. Board Pre-print CD-ROM*, (Mira Digital Publishing, Washington, D.C).

- [70] Hough, P.V.C. (1962) Method and means for recognizing complex patterns, U.S. patent 3 069 654.
- [71] Hue, C. Le Cadre, J.-P. Pérez, P. (2001) A particle filter to track multiple objects, *In: WOMOT '01: Proceedings of the IEEE Workshop on Multi-Object Tracking (WOMOT'01)*, IEEE Computer Society, Washington, DC, USA, 2001.
- [72] Hughes, R. L. (2003) The flow of human crowds, *Annual Review of Fluid Mechanics* **35**, 169.
- [73] Intille, S., Davis, J., and Bobick, A. (1997) Real-time closed-world tracking, *In: Proc. Conf. Computer Vision and Pattern Recognition..*
- [74] Isard, M. A. (1998) Visual Motion Analysis by Probabilistic Propagation of Conditional Density, PhD thesis, University of Oxford.
- [75] Jebara, T., Russell, K., and Pentland, A. (1998) Mixtures of eigenfeatures for real-time structure from texture, *In: Proc. 6th Int. Conf. on Computer Vision*, 128–135. Mumbai, India.
- [76] Johansson, A. (2004) Pedestrian simulations with the social force model, Master's thesis, Dresden University of Technology.
- [77] Johansson, A., Chan, S., Hamdar, S., Al-Bosta, S., Haase, K., and Helbing, D. (2008) Towards automatic analysis and forecasting of pilgrim flow by video analysis and simulation, *Specialized Architectural Engineering Magazine issued by Ministry of Municipal and Rural Affairs, Kingdom of Saudi Arabia*.
- [78] Johansson, A. and Helbing, D. (2007) Pedestrian flow optimization with a genetic algorithm based in Boolean grids. Pages 237-272 in: N. Waldau, P. Gattermann, H. Knoflacher and M. Schreckenberg (eds.) *Pedestrian and Evacuation Dynamics 2005* (Springer, Berlin).
- [79] Johansson, A. and Helbing, D. (2006) Crowd dynamics. Pages 449–471 in: B.K.Chakrabarti, A.Chakraborti, A. Chatterjee (eds.) *Econophysics and Socio-physics – Trends and Perspectives* (Wiley, Weinheim).
- [80] Johansson, A., Helbing, D. and Shukla, P. K. (2007) Specification of the social force pedestrian model by evolutionary adjustment to video tracking data, *Advances in Complex Systems* **10**, 271–288.

- [81] Johansson, A., Helbing, D., Al-Abideen, H. Z., and Al-Bosta, S. (2008) From Crowd dynamics to crowd safety: A video-based analysis *Advances in Complex Systems* **11** (4), 497–527.
- [82] Kass, M., Witkin, A., and Terzopoulos, D. (1987) Snakes: Active contour models, *In: Proc. 1st Int. Conf. on Computer Vision*, 259–268.
- [83] Kerridge, J., Keller, S., Chamberlain, T., and Sumpter, N. (2007) Collecting pedestrian trajectory data in real-time, Pages 27–39 in: N. Waldau, P. Gattermann, H. Knoflacher and M. Schreckenberg (eds.) *Pedestrian and Evacuation Dynamics 2005* (Springer, Berlin).
- [84] Klüpfel, H. (2003) A Cellular Automaton Model for Crowd Movement and Egress Simulation, PhD thesis, Universität Duisburg-Essen.
- [85] Klüpfel, H. (2007) The simulation of crowd dynamics at very large events – calibration, empirical data, and validation. Pages 285–296 in: N. Waldau, P. Gattermann, H. Knoflacher and M. Schreckenberg (eds.) *Pedestrian and Evacuation Dynamics 2005* (Springer, Berlin).
- [86] Koss, L. L. and Brumley, A .T. (2003) in: *Pedestrian and Evacuation Dynamics, E. R. Galea (ed.)*, 257–265, CMS Press, London.
- [87] Kretz, T., Grünebohm, A. and Schreckenberg, M. (2006) Experimental study of pedestrian flow through a bottleneck, *J. Stat. Mech* P10014.
- [88] Kretz, T. (2007) Pedestrian Traffic – Simulation and Experiments, PhD thesis, Universität Duisburg-Essen.
- [89] Koenderink, J. and van Doorn, A. (1975) Invariant properties of the motion parallax field due to the movement of rigid bodies relative to an observer, *Optica Acta* **22** (9), 773–791.
- [90] Le Bon, G. (1896) The Crowd: A Study of the Popular Mind. (The Macmillan Co., New York).
- [91] LeCun, Y., Huang, F.-J. and Bottou, L. (2004) Learning Methods for Generic Object Recognition with Invariance to Pose and Lighting, *Proceedings of CVPR'04*, (IEEE Press).
- [92] Lewin, K. (1951) Field Theory in Social Science, (Harper & Brothers, New York).

- [93] Lipton, A. J., Fujiyoshi, H. and Patil, R. S. (1998) Moving target classification and tracking from real-time video, In: *Proc. IEEE Image Understanding Workshop*, 129–136.
- [94] Lövås, G. G. (1994) *Transportation Research* **28B**, 429.
- [95] Meyer-König, T., Klüpfel H., and Schreckenberg, M. (2002) in: Assessment and analysis of evacuation processes on passenger ships by microscopic simulation, *Pedestrian and Evacuation Dynamics*, M. Schreckenberg and S. D. Sharma (eds.), Springer, Berlin, 297–302.
- [96] Michalewicz, Z., Fogel, D. B. (2000) How to Solve It: Modern Heuristics (Springer-Verlag).
- [97] Mintz, A. (1951) Non-adaptive group behavior. *The Journal of Abnormal and Normal Social Psychology* **46**, 150–159.
- [98] Mohamed, M. N. (1996) Hajj & Umrah From A to Z (Mamdouh Mohamed).
- [99] Molnár, P. (1996) Modellierung und Simulation der Dynamik von Fußgängerströmen, PhD thesis, University of Stuttgart.
- [100] Mori, M. and Tsukaguchi, H. (1987) A new method for evaluation of level of service in pedestrian facilities, *Transportation Research A* **21**(3), 223–234.
- [101] Musse, S. R., Babski, C., Capin, T., and Thalmann, D. (1998) Crowd modelling in collaborative virtual environments, Pages 115–123 in: *Proceedings of the 1998 ACM symposium on Virtual reality software and technology*.
- [102] Nagai, R., Fukamachi, M. and Nagatani, T. (2006) Evacuation of crawlers and walkers from corridor through an exit *Physica A* **367**, 449–460.
- [103] Nilsson, D. and Johansson, A. (2008) Social influence during the initial phase of a fire evacuation - Analysis of evacuation experiments in a cinema theatre, *Fire Safety Journal*, in print.
- [104] O’Sullivan, D., and Haklay, M. (2000) Agent Based Models and Individualism: Is the World Agent-Based? *Environment and Planning A* **32** 1409–1425.
- [105] Okazaki, S. (1979) A study of pedestrian movement in architectural space, Part 1: pedestrian movement by the application on of magnetic models, *Trans. of A.I.J..*

- [106] Osadchy, M., LeCun, Y. and Miller, M. (2007) Synergistic face detection and pose estimation with energy-based models. *Journal of Machine Learning Research* **8**, 1197–1215.
- [107] Papageorgiou, C. and Poggio, T. (1999) Trainable pedestrian detection, In: *Proceedings of the International Conference on Image Processing (ICIP 99)*.
- [108] Penn, A. and Turner, A. (2001) Space syntax based agent simulation, In: *1st International Conference on Pedestrian and Evacuation Dynamics, 2001* (University of Duisburg, Germany).
- [109] Peters, K., Johansson, A., Helbing, D. (2005) *Künstliche Intelligenz* **4**, 11.
- [110] Pheasant, S. T. (1998) Bodyspace: Anthropometry, Ergonomics, and the Design of Work (Taylor and Francis, London).
- [111] Philomin, V., Duraiswami, R. and Davis, L. (2000) Pedestrian tracking from a moving vehicle, *In: Procs. IEEE Intelligent Vehicles Symposium*.
- [112] Polus, A., Schofer, J. L. and Ushpiz, A. (1983) Pedestrian flow and level of service, *Journal of Transportation Engineering* **109**, 46–56.
- [113] Popovic, J., Seitz, S. M., and Erdmann, M. (2003). Motion sketching for control of rigid-body simulations, *ACM Transactions on Graphics* **22** (4), (SIGGRAPH '03 Conference Proceedings), 1034–1054.
- [114] Predtetschenski, W. M. and Milinski, A. I. (1971) Personenströme in Gebäuden - Berechnungsmethoden für die Projektierung (Rudolf Müller, Köln-Braunsfeld)
- [115] Reynolds, C. W. (1987) Flocks, herds, and schools: A distributed behavioral model, *Computer Graphics* **21**(4), (SIGGRAPH '87 Conference Proceedings), 25–34.
- [116] RiMEA: www.rimea.de
- [117] Seyfried, A., Steffen, B., Klingsch, W., Boltes, M. (2005) The fundamental diagram of pedestrian movement revisited, *J. Stat. Mech.*, P10002.
- [118] Seyfried, A., Steffen, B., Winkens, A., Rupprecht, T., Boltes, M. and Klingsch, W. (2007) Empirical data for pedestrian flow through bottlenecks in: Traffic and Granular Flow 2007, Proceedings of the International Conference, Jun. 20-22 2007, Paris Orsay, France, Springer (Berlin), 2007 (to appear).

- [119] Shari, A. (2005) HAJJ: Reflection on Its Rituals (Islamic Publications International).
- [120] Shukla, P. K. (2005) Modeling and Simulation of Pedestrians, Master's thesis, Dresden University of Technology.
- [121] Sobel, I. and Feldman, G. (1968) A 3x3 Isotropic Gradient Operator for Image Processing, presented at a talk at the Stanford Artificial Project in 1968, unpublished.
- [122] Smith, R. A. and Dickie, J. F. (eds.) (1993) Engineering for Crowd Safety (Elsevier, Amsterdam).
- [123] Sorenson, H. and Alspach, D. (1971) Recursive Bayesian estimation using Gaussian sums, *Automatica*, **7**, 465–479.
- [124] Steffen, B and Seyfried, A. (2008) The repulsive force in continuous space models of pedestrian movement. *arXiv:0803.1319v1 [physics.soc-ph]*.
- [125] Still, K. (2000) Crowd Dynamics, PhD thesis, University of Warwick.
- [126] Teknomo, K. (2002) Microscopic Pedestrian Flow Characteristics: Development of an Image Processing Data Collection and Simulation Model, PhD thesis, Tohoku University.
- [127] Teknomo, K., Takeyama, Y. and H. Inamura (2001) Tracking System to Automate Data Collection of Microscopic Pedestrian Traffic Flow, *In: Proceeding of The 4th Eastern Asia Society For Transportation Studies* **3**:1, 11–25.
- [128] Thompson, P. (2004) Simulex: Simulated people have needs too. www.iesve.com, August 2004.
- [129] Torr, P. and Murray, D. (1994) Stochastic motion clustering, Pages 328-337 in: *Proc. 3rd European Conf. Computer Vision* (Springer-Verlag).
- [130] Tilch, B. and Helbing, D. (2000) Evaluation of single vehicle data in dependence of the vehicle-type, lane, and site, Pages 333–338 in: *Traffic and Granular Flow '99: Social, Traffic, and Granular Dynamics*, eds. D. Helbing, *et al.*, (Springer, Berlin, 2000).

- [131] Treiber, M., Kesting, A., and Helbing, D. (2006) Delays, inaccuracies and anticipation in microscopic traffic models, *Physica A* **360**, 71–88.
- [132] Treuille, A., Cooper, S., and Popovic, Z. (2006) Continuum crowds, *ACM Transactions on Graphics* **25** (3), 1160–1168.
- [133] Treuille, A., Lee, Y., and Popovic, Z. (2007) Near-optimal character animation with continuous control, *ACM Transactions on Graphics* **26** (3), (SIGGRAPH '07 Conference Proceedings).
- [134] Velastin, S. A. Yin, J. H. Davies, A. C. Vicencio-Silva, M. A. Allsop, R. E. and Penn, A. (1994) Automated measurement of crowd density and motion using image processing, Pages 127–132 in: *Seventh International Conference on Road Traffic Monitoring and Control (CP391)*
- [135] Watts, J. M. (1987) *Fire Safety Journal* **12** 237.
- [136] Weidmann, U. (1993) Transporttechnik der Fußgänger. *ETH-Zürich, Schriftenreihe IVT-Berichte* **90**, Zürich.
- [137] Welford, A. T. (1980) Choice reaction time: Basic concepts. In A. T. Welford (Ed.), *Reaction Times*. Academic Press, New York, pp. 73–128.
- [138] Wiseman, R. (2007) Quirkology, pace of life project, <http://www.paceoflife.co.uk>.
- [139] Woodworth, R. S. and Schlosberg, H. (1954) *Experimental Psychology*. Henry Holt, New York.
- [140] Wren, C., Azarbayejani, A., Darrell, T., and Pentland, A. (1997) Pfnder: Real-time tracking of the human body, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **19** (7), 780–785.
- [141] Xu, F., Liu, X., Fujimura, K. (2005) Pedestrian detection and tracking with night vision, *ITS(6)* **1**, 63–71.
- [142] Yu, W. J., Chen, R., Dong, L. Y., and Dai, S. Q. (2005) Centrifugal force model for pedestrian dynamics, *Phys. Rev. E* **72**, 026112.
- [143] Yu, W. and Johansson, A. (2007) Modeling crowd turbulence by many-particle simulations, *Phys. Rev. E* **76**, 046105

- [144] Zhao, H. and Shibasaki, R. (2005) A novel system for tracking pedestrians using multiple single-row laser-range scanners. *IEEE Transactions on Systems, Man and Cybernetics, Part A* **35** (2), 283–291.
- [145] Zhu, W. and Timmermans, H. (2005) Exploring Pedestrian Shopping Decision Processes, Pages 145–154 in: N. Waldau, P. Gattermann, H. Knoflacher and M. Schreckenberg (eds.) *Pedestrian and Evacuation Dynamics* (Springer, Berlin).