

Tracking Individual Targets in High Density Crowd Scenes Analysis of a Video Recording in Hajj 2009

Mohamed H. Dridi

Institute of Theoretical Physics 1, University of Stuttgart, Stuttgart, Germany

Email: Mohamed.dridi@itp1.uni-stuttgart.de

Received 29 January 2015; accepted 12 February 2015; published 30 March 2015

Copyright © 2015 by author and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

In this paper we present a number of methods (manual, semi-automatic and automatic) for tracking individual targets in high density crowd scenes where thousands of people are gathered. The necessary data about the motion of individuals and a lot of other physical information can be extracted from consecutive image sequences in different ways, including optical flow and block motion estimation. One of the famous methods for tracking moving objects is the block matching method. This way to estimate subject motion requires the specification of a comparison window which determines the scale of the estimate. In this work we present a real-time method for pedestrian recognition and tracking in sequences of high resolution images obtained by a stationary (high definition) camera located in different places on the Haram mosque in Mecca. The objective is to estimate pedestrian velocities as a function of the local density. The resulting data of tracking moving pedestrians based on video sequences are presented in the following section. Through the evaluated system the spatio-temporal coordinates of each pedestrian during the Tawaf ritual are established. The pilgrim velocities as function of the local densities in the Mataf area (Haram Mosque Mecca) are illustrated and very precisely documented. Tracking in such places where pedestrian density reaches 7 to 8 persons/m² is extremely challenging due to the small number of pixels on the target, appearance ambiguity resulting from the dense packing, and severe inter-object occlusions. The tracking method which is outlined in this paper overcomes these challenges by using a virtual camera which is matched in position, rotation and focal length to the original camera in such a way that the features of the 3D-model match the feature position of the filmed mosque. In this model an individual feature has to be identified by eye, where contrast is a criterion. We do know that the pilgrims walk on a plane, and after matching the camera we also have the height of the plane in 3D-space from our 3D-model. A point object is placed at the position of a selected pedestrian. During the animation we set multiple animation-keys (approximately every 25 to 50 frames which equals 1 to 2 seconds) for the position, such that the position of the point and the pedestrian overlay nearly at every time. By combining all these variables with the availa-

ble appearance information, we are able to track individual targets in high density crowds.

Keywords

Pedestrian Dynamics, Crowd Management, Crowd Control, Objects Tracking

1. Introduction

Crowd simulation has found its way into computer science, computer visualizations and the computer simulation of oriented building construction and crowd management (Predtetschenski & Milinski, Russ: 1969, Germ: 1971). With continuously growing population around the world and with enormous evolution in the different modes of transportation in the last decade, a lot of paper have appeared with increasing interest in modelling crowd and evacuation dynamics. Thus the simulation of pedestrian flows has become an important research area. Pedestrian models are based on macroscopic or microscopic behaviour.

The evolution and design of any pedestrian simulation model requires a lot of information and data.

A number of variables and attributes arises from empirical data collection and need to be considered to develop and calibrate a (microscopic) pedestrian simulation model.

For this reason we use different tools and develop different methods to collect the microscopic data and to analyse microscopic pedestrian flow. It is very important to mention that the pedestrian data collection especially in a dangerous situation is still very much in its infancy. An aim of this study is to establish more clearness and understanding about the microscopic pedestrian flow characteristics. Manual, semi manual and automatic image processing data collection systems are developed. Many published studies show that the microscopic speed obey a normal distribution with a mean of 1.38 m/second and a standard deviation of 0.37 m/second. The acceleration distribution also resemblances a normal distribution with an average of 0.68 m/square second (Knoblauch et al., 1996; Fruin et al., 1970; O'Flaherty, 1996).

For the evolution and development of pedestrian microscopic simulation models, a lot of data was collected with the help of video recording and tracking of moving entities in the pedestrian flow using the coordinates of the head path that was established through image processing. A large trajectory dataset has been restored. For the observation of pedestrian flows in public places a Sony camera was used. This observation was in different places where the pilgrims performed their rituals. Many variables can be gathered to describe the behaviour of pedestrians from different points of view. This paper describes how to obtain variables from video taking and simple image processing that can represent the movement of pedestrians (pilgrims) and its variables. Moreover in this work we try to understand several parameters influencing the pedestrian behaviour in riots or panic situations.

For obtaining empirical data different methods were used, including automatic and manual methods. We have analysed video recordings of the crowd movement in the Tawaf in Mosque/Mecca during the Hajj on the 27th of November, 2009. We have evaluated unique video recordings of a 105×154 m large Mataf area taken from the roof of the Mosque, where upto 3 million Muslims perform the Tawaf and Sa'y rituals within 24 hours.

Both Microscopic Video Data Collection and Microscopic Pedestrian Simulation Model generate a database called PedFlow database. The properties and characteristics that are capable of explaining microscopic pedestrian flow are illustrated. A comparison between average instantaneous speed distributions describing the real world is obtained from different methods, and how they can be used in the calibration and validation of the simulation tools are explained.

2. Related Work

Typically, manual counting is performed by tally sheet or mechanical or electronic count board to collect density and speed data for pedestrian. Pedestrian behaviour studies are collected by manual observation or video recording in different public places like corridors side walks and cross walks. The effectiveness of the data (pedestrian speed) collected on any observed area is strongly related to the number of pedestrians in the flow. The relationship between speed, flow, and pedestrian density for a crowd population or human group has been

published in many fundamental diagrams developed by Fruin (Fruin, 1987) and others (Chattaraj et al., 2009). Though for many reasons the method has been used to detect and count vehicles in automatic way, it cannot be used to detect pedestrians, since this system has been evaluated through pneumatic tube or inductance loops. As we can deduce from later work on this technology, the possibility of applying this method to reproduce trajectory and motion prediction is still in a discussion phase.

Other approaches use a neural network framework recursively to predict pedestrian motion and trajectory (Bulpitt & Sumpter, 1998). However, the pedestrian trajectories in this system are calculated with incorrect simplifications. In particular, only the nearest neighbour trajectories are considered. The main shortcoming of such an estimation is that there is no uncertainty in this prediction, moreover a comparison of different path prediction shows this is still far from the reality in order to predict that all objects will follow the same set of paths exactly.

A method which allowed people to count based on video texture synthesis and to reproduce motion in a novel way was introduced by Heisele and Woehler (Heisele & Woehler, 1998). The method works under the assumption that people can be segmented from the moving background by means of appearance or motion properties. The scene image is clustered based on the color and position (R, G, B, X, Y) of pixel. The appearance of each pixel in a video frame is modelled as a mixture of Gaussian distributions. A algorithm is used that matches a spherical crust template to the foreground regions of the depth map. Matching is done by a time delay neural network for object recognition and motion analysis.

A significant task in video intelligence systems is the extraction of information about a moving objects, e.g. detecting a moving crowd with PedCount (a pedestrian counter system using CCTV) was developed by Tsuchikawa (Tsuchikawa et al., 1995). It extracts the object using the one line path in the image by background subtraction to make a space-time (X-T) binary image. The direction of each travelling pedestrian is realized by the attitude of pedestrian region in the X-T image. They reported the need of background image reconstruction due to image illumination change. An algorithm to distinguish moving object from illumination change is explained based on the variance of the pixel value and frame difference.

3. Analyse of the Video Taking in Hajj2009

3.1. Introduction

The electronic and digital revolution in video techniques during recent years has made it possible to gather detailed data concerning pedestrian behaviour, both in experiments and in real life situations (Hoogendoorn & Daamen, 2005, Helbing et al., 2008; Boltes et al., 2010). The big challenge is to develop a new efficient method of defining and measuring basic quantities like density, flow and speed. Basic quantities of pedestrian dynamics are the density ρ [1/m²] in an area A and the velocity v [m/s] of persons or a group of persons, and the flow through a door or across a specific line $Q = v(r, t)\rho(r, t)$ [1/s]. The measurements also yield mean values of these quantities. The task is to improve the given methods such that they allow going fairly close to the real data of the crowd quantities. The methods presented here are based on video tracking of the head from above. Note that tracking of e.g. a shoulder or the chest might be even better, though more difficult to obtain.

The density distribution knowledge in a very crowded area allows us to draw a so-called density map to show us congestion directly as regions of high density. The relationship between the pedestrian density ρ and the pedestrian maximum walking speed v_{\max} are formalized into a graph known as the fundamental diagram $v_{\max} = f(\rho)$ (Fruin et al., 1970). Since pedestrians move slower in a region of high density, the simulated particles should update their speed with the surrounding circumstances to maximize their rate of progress towards their goals.

3.2. Data Collection and Type of Observation

Tawaf observations at the Haram mosque in Mecca were made during Hajj 2009 by Mr. Faruk Oksay. The Mataf area has 10 entrances/exits. The flow of the Tawaf is controlled. All pilgrims begin and end their Tawaf at the same place (Figure 1). The number of pilgrims during this period is sufficient to observe the behaviour of high density crowd dynamics.

Figure 1 shows the main gate doors, side entrances, stairs to the Mataf open air of the Haram.

All observations took place on Friday, 27th November 2009, corresponding to 10th of Dhu al-Hijjah 1430

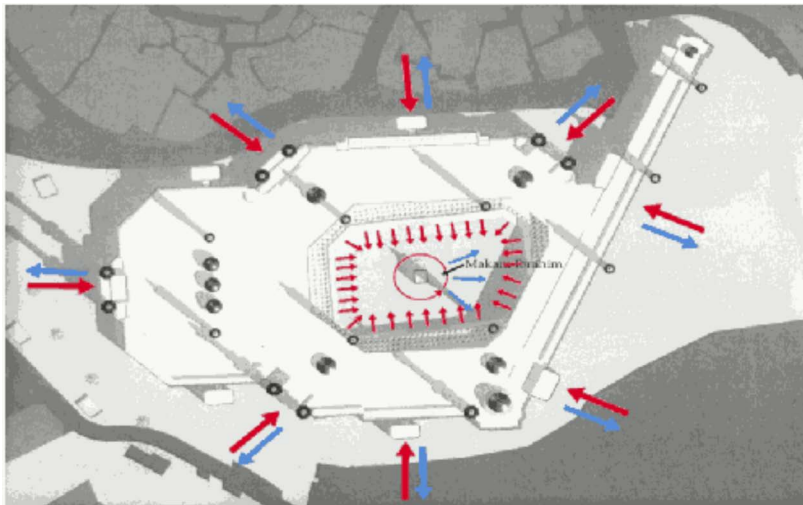


Figure 1. Overview of the Mosque with entrance doors.

Hijri in the afternoon. During the total observation period of three hours, three prayers (Midday, Asr and Maghreb (sunset-prayer)) were performed, where in this time the Mataf area came to a standstill (**Figure 2**). Our video observations showed that the pilgrim had the desire to be near the Kaaba. Therefore approximately 70 percent (visually detected on video) of the pilgrims performed their Tawaf movement near the Kaaba wall, which caused a high density in this area.

In **Figure 2**, one can see all of the pilgrims perform the prayer ritual in the holy mosque in Mecca.

The Tawaf around the Kaaba is a periodic movement for the time between two prayers. The observed number of pilgrims performing their Tawaf ritual at the Mataf area increases slowly after every prayer until the Mataf attains its maximum capacity (**Figure 3**).

Figure 3 shows a typical pedestrian movement in the Mataf area over daytime. During prayer times individuals stand still and therefore movement equals approximately zero. The fluctuations in the velocity flow are created by the turbulence in the pedestrian flux. Note that the average local density in a specific location in the Mataf area exceeds 8 persons/m² during the Hajj periods (**Figure 6** and **Figure 7**).

3.3. Goals

Our first goal is to identify new methods and create a test system capable of extracting pedestrian movement information from video, similar to that collected by our HD-Cameras in the Hajj-2009, such that any movement can be analysed to spot suspicious activity. This task to collect pedestrian data and extract pedestrian motion from video sequences requires an involvement and development of appropriate methods, followed by further analysis of this data to identify emergent motion or crossing trajectories.

The secondary goal is to identify the limitations of the approach including the system and data requirements for the techniques to work more effectively. More specific, the project goals are:

- To develop a framework for video and image analysis,
- To develop an approach and relevant diagnostic software to collect movement data from video,
- To identify the requirements for such methods to work effectively, such as image quality, resolution and orientation,
- To identify how to interpret movement information,
- To interpret the movement data and examine abnormal behaviour,
- To design and produce a working implementation that demonstrates the above goals,
- To identify approaches that could further improve the system.

4. Estimation of Crowd Density

There are different techniques developed to extract information describing the position of pedestrians in a location, but not all of them are appropriate for detecting and pursuing pedestrian movement under different

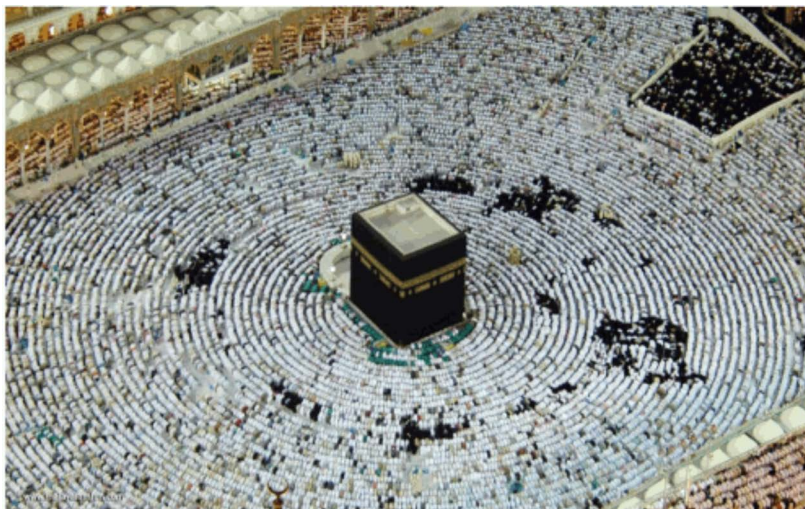


Figure 2. Pilgrims performing the prayer ritual in the Mosque in Mecca. During performance of prayer the Tawaf come to standstill, there are no movement around the Kaaba.

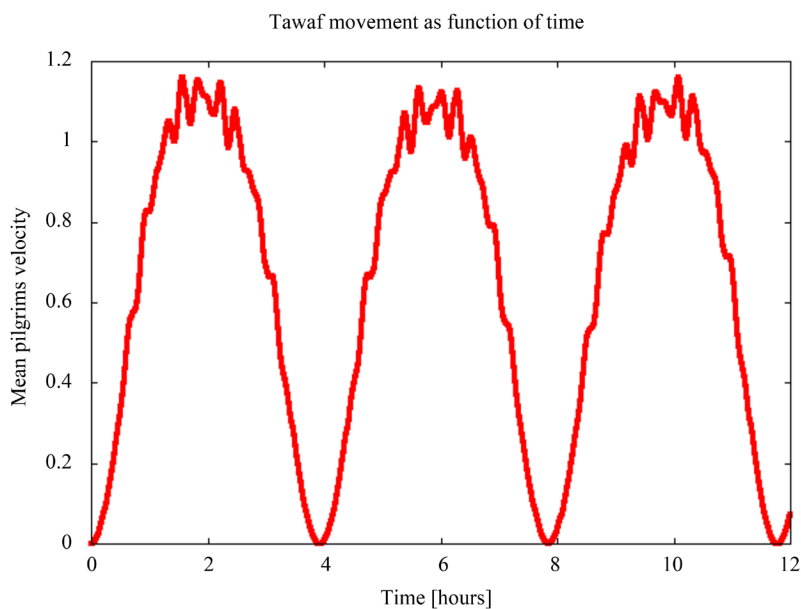


Figure 3. Mean velocity of pilgrims in the Mataf area as a function of time over 12 hours.

extreme weather conditions. In their published work (Papageorgiou & Poggio, 1999), Papageorgiou and Poggio developed a system attempting to recognize human figures based on pixel similarities through a large training set of figures under various light and weather conditions. To identify the movement of the figures, the system analyses the similarity between matches of consecutive frames. This method works quite well when the training set is large, but requires a high computational efficiency which achieves processing rates of 10 Hz (Papageorgiou & Poggio, 1999). The study shows that accurate recognition can be done with coarse image data.

Another approach to estimate crowd density is based on texture analysis. Velastin et al. (Marana et al., 1998) assumed that crowds with high density possessed texture properties. The proposed method, texture features were computed for the whole image and applied to crowd density estimation (Zhang & Li, 2012). In particular, all displayed textures, like wavelets (Verona & Marana, 2001; Li et al., 2006) and the gray level dependence matrix (Wu, 2006; Sen et al., 2009), were used to estimate crowd density. The results exhibited, how effective

statistical analysis of texture display was compared with neural networks when measuring crowd density. Unfortunately, this system examined only static images and cannot cover crowd motion, but the techniques can be used to track pedestrian movements.

Other strategies based on image segmentation were pursued by Heisele and Woehler (Heisele & Woehler, 1998), where raw data was filtered to split the image into segments, which were then analysed. Those images that matched particular shapes were analysed further. This approach allowed to distinguish different images with common color and luminescence.

4.1. Data

The required data on pedestrian behaviour (e.g. density-effect, shock-waves-effect,...) in the Haram can be done from our video recordings. All observed effects can be analysed by simply watching the recorded videos. But if we want to extract data like walking speeds from such observations we have to examine the videos frame by frame. This is very time consuming. As a result of this, and the need for more efficient data, the idea arose to use an automatic detection system. At that time no sufficient system was available for the detection of human bodies, therefore some essential requirements were formulated. From the requirements we derived an idea to formulate an image processing system with the help of other programs, such as Optical Flow with OpenCV (<http://opencv.org/>) and Quest3D (<http://www.quest3d.com/>). The materials used for this test were videos recorded at an outdoor piazza of the Haram mosque in Mecca where people congregated at different times during one day, simulating a surveillance application. The data content had a wide range of crowd densities, from very low to very high. Three different datasets, labelled morning observation, afternoon observation and combined observation (before and after the prayer times) were used. Each data set had 20 selected images with high resolution. Examples of images are shown in Figure 4 and Figure 12.

In order to collect pedestrian data and to study pedestrian traffic flow operations on a platform in detail, observations were also made from a platform of the Haram mosque in Mecca. These observations concerned pilgrim walking speeds and density distributions on the Mataf area and (individual) walking times as functions of the distance from the Kaaba wall.

4.2. Methods

4.2.1. Manual Estimation of Crowd Density

The estimation of crowd density is an important criterion for the validation of our simulation tools. Processing is done in three levels.

- Existing footage is loaded on a 3D program as a backplate;
- From several provided 2D-architectural drawings we build a 3D model of the mosque;
- A virtual camera has to be matched in position, rotation and focal length to the original camera so that the features of the 3D-model match the features of the filmed mosque. As the dimensions of the mosque are known, we then establish a grid of regular cells on the Mataf area, each one of which has a size of $5\text{ m} \times 5\text{ m}$ (see Figure 5). Through image editing software, we start a manual counting process. This regular grid is used to observe the density behaviour over all of the Mataf area, from the nearest range to the Kaaba wall up to outside of the Mataf and the accumulation process (by the Black Stone and Maqam Ibrahim). The results of this investigation are shown in Figures 6(a)-(d) and illustrate the behaviour of the pilgrim density on the Mataf area at different times during the day.

With a new computer algorithm developed within this investigation, where the Mataf area is divided in regular cells. The number of pedestrians in every cell as function of time is determined through repeating the counting process many times. The average value is identified as local density $\rho(\mathbf{r}, t)$. The data extracted from the videos allowed us to determine not only densities in larger areas, but also local densities, speeds and flows. As an example the density distribution on the Mataf area is shown in Figure 6. The data was obtained by semi-manual evaluation.

Dependence of the density distribution on the Mataf as function of time. Figure 6 shows density decline curves for different distances from the Kabaa in a specific time. The curves indicate that the local density amount vary strongly over the ($0 < x < 40\text{ m}$) range.

Figure 7 shows the pedestrian density distribution on the Mataf area as a function of the position \mathbf{r} and

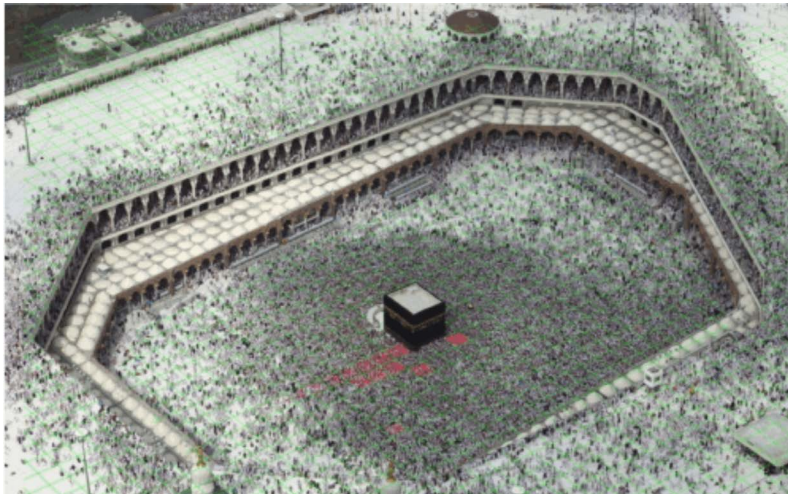


Figure 4. The Haram piazza top view (observation area).

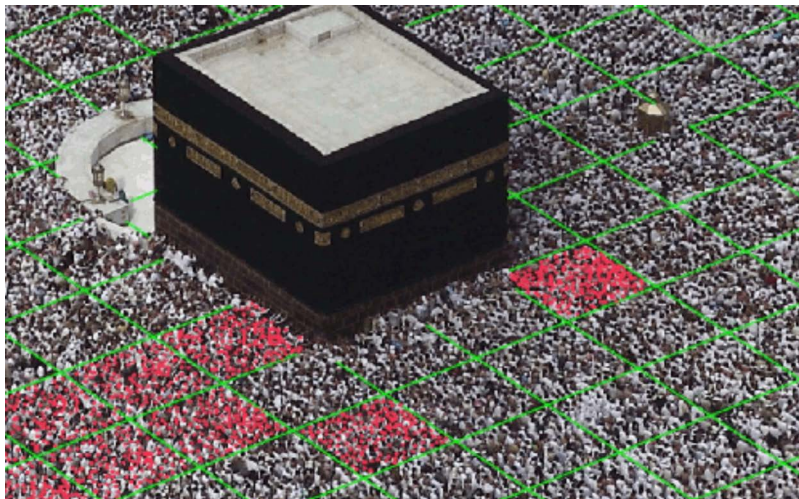


Figure 5. Grid of regular cells with dimension of $5\text{ m} \times 5\text{ m}$.

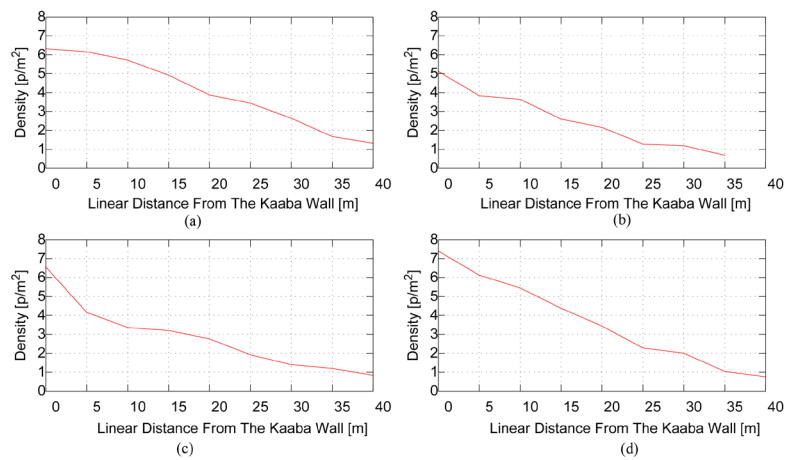


Figure 6. Decrease of pedestrian density on the Mataf Area as function of the distance from the Kaaba wall: (a) before Mid-Day prayer; (b) shortly after Mid-Day Prayer; (c) half-hour after Mid-Day Prayer; (d) Rush Hour.

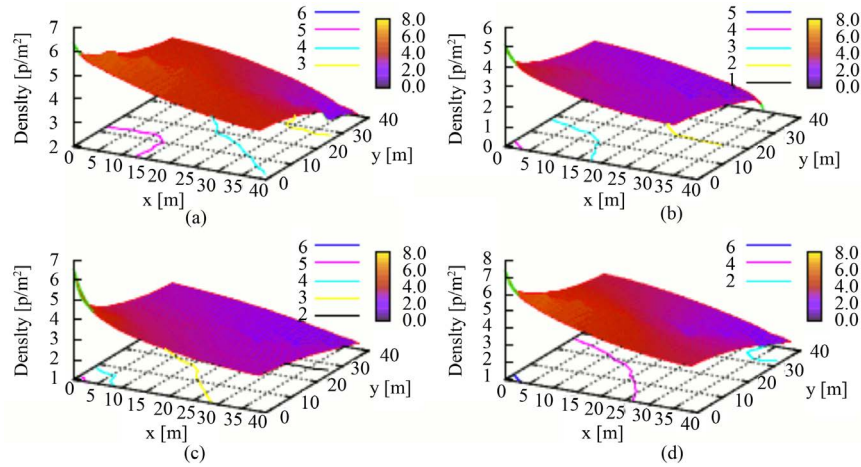


Figure 7. The Density Distribution on the Mataf Area $\rho(\mathbf{r}, t)$ with $\mathbf{r} = (x, y)$. The figure also shows the density index (persons/m²). (a) before Mid-Day prayer ($t = t_{\text{before Mid-Day}}$); (b) shortly after Mid-Day Prayer ($t = t_{\text{shortly after Mid-Day}}$); (c) half-hour after Mid-Day Prayer ($t = t_{\text{half-hour after Mid-Day}}$); (d) Rush Hour ($t = t_{\text{Rush}}$).

time t . One clearly recognizes density waves with maximum density package near the Kaaba wall. There the average local density can reach a critical value of 7 to 8 persons/m². The congested area increases the local density to a critical and dangerous amount. As a consequence the pedestrians begin to push to increase their personal space and create shock-waves propagating through the crowd, which can be seen as density waves, or density packages.

The Density map illustrates how the pedestrian density decreases from the inside to outside of the Mataf area, (see [Figure 8](#)). As we have mentioned that in the Mataf area pedestrians move in the restricted space, the layout is gradually painted in different colors. The color of every point of the space corresponds to the current density in this particular area. The density map is constantly repainted according to the actual values: when the density changes in some point, the color changes dynamically to reflect this change. In case of zero density the area is not painted at all (see [Figures 8\(a\)-\(d\)](#)).

During the rush hour in a Hajj period the local density in the Mataf area reaches the maximum as we can see in the following [Figure 7\(d\)](#) and [Figure 8\(d\)](#). The local density can reach 8 to 9 persons/m² in a specific time during the day. The maximal density concentrates near the Kaaba wall.

Densities over time and space. We observed the density behaviour on the Mataf area at different times during the day, before and after the prayer, and we compared this density with the simulation density results. The maximum registered density was 7 to 8 persons/m² and this represents a high crowd density. The results of the estimation based on the statistical method, presented in [Figures 6-8](#), reached a mean of 92 percent correct estimations. It was possible to verify that the results were quite good for all evaluated images except for the one made up of high density crowd images, which reached only 84 percent correct estimations. In the Mataf area, near the black stone, the pilgrim density reached over 9 persons/m². For this reason it was very difficult to recognize and track every head and as a result, a 100 percent correct estimation would be very difficult. All statistical results illustrating the density distribution at the Mataf area at different time intervals are demonstrated in the [Figure 9](#).

4.2.2. Automatic Estimation of Crowd Density

This part of the dissertation considers the role of automatic estimations of crowd density and their importance for the automatic monitoring of areas where crowds are expected to be present. A new technique is proposed which is able to estimate densities ranging from very low to very high concentrations of people. This technique is based on the differences of texture muster on the images of crowds. Images of low density crowds exhibits rough textures, while images with high densities tend to present finer textures. The image pixels are classified in different texture classes, and statistics of such classes are used to estimate the number of people. The texture classification and the crowd density estimation are based on self-organizing neural networks. Results obtained to estimate the number of people in a specific area of the Haram Mosque in Mecca are presented in [Figure 10](#).

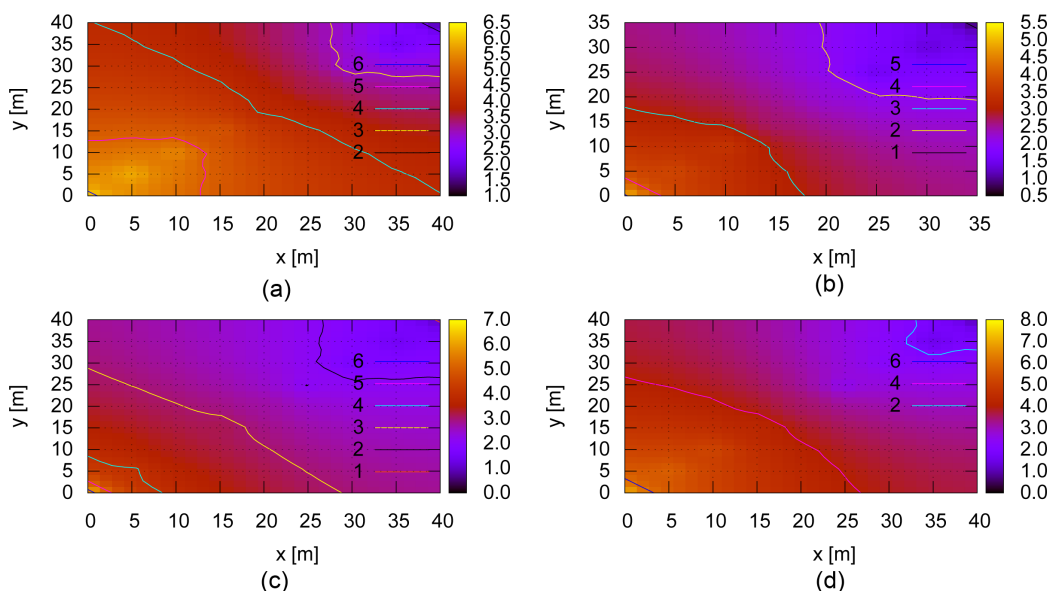


Figure 8. The Density map indicates that the highest pedestrian density in the area of the Kaaba: (a) before Mid-Day prayer; (b) shortly after Mid-Day Prayer; (c) half-hour after Mid-Day Prayer; (d) Rush Hour.

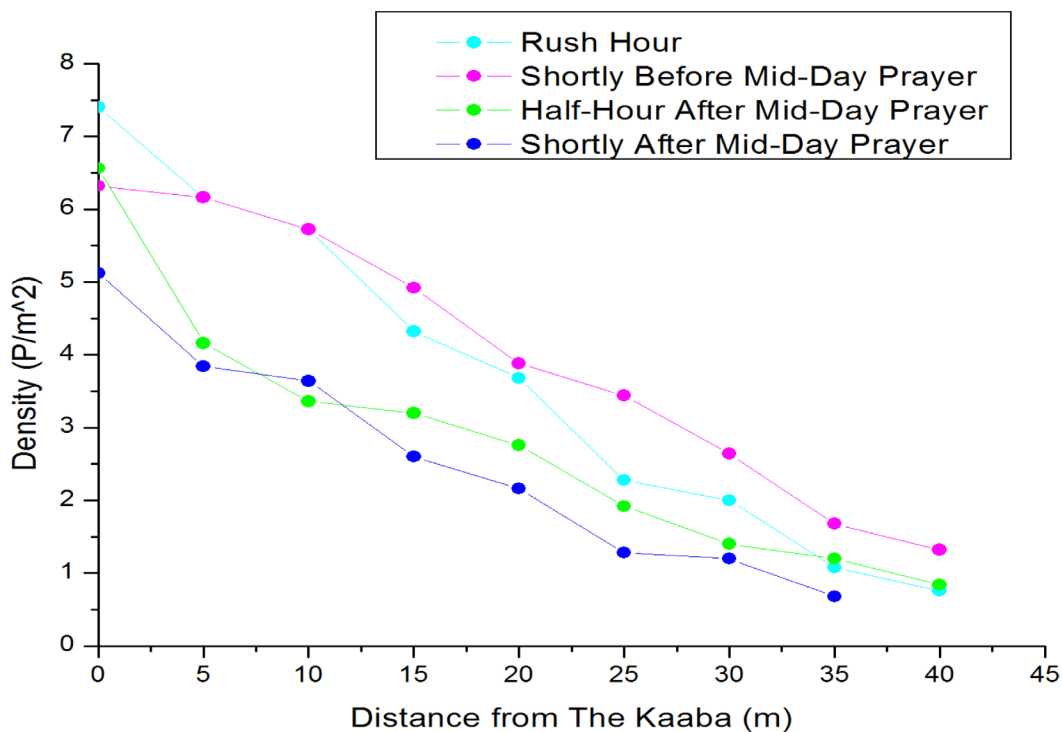


Figure 9. Crowd density on the Mataf area in different time intervals. Highest density in the area of the Kaaba.

4.3. Data Analysis

In the latter paragraphs we focus on crowd density estimation for several reasons. According to the crowd disasters study by Helbing, Al-Abideen and Johansson (Helbing et al., 2007), one of the most important aspects to keep a crowd safe is to predict and identify areas with high density crowds preventing large crowd pressures to be built up. Areas where crowds are likely to build up should be identified prior to the event or operation of

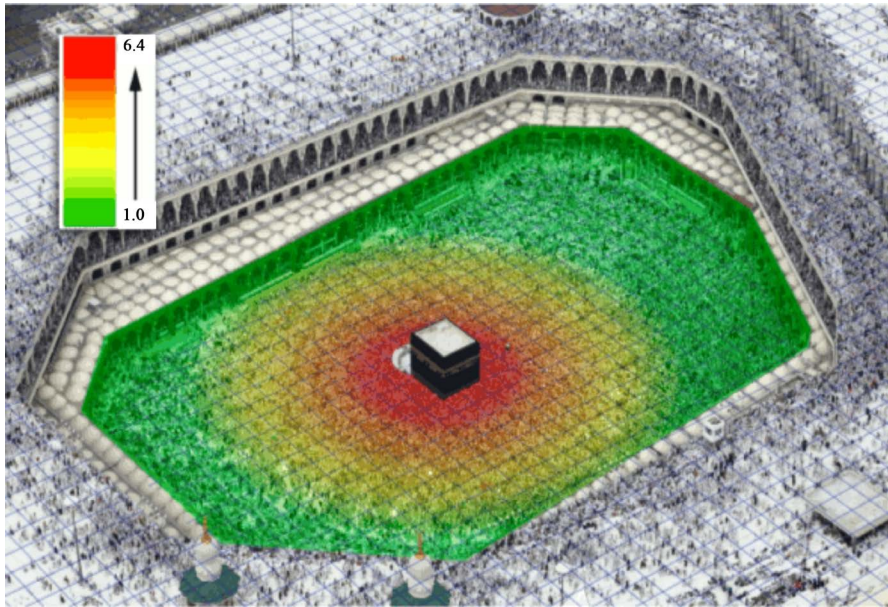


Figure 10. Density distribution in the Mataf area before Mid-Day prayer. Red color indicates high density, where green color indicates low pilgrims density.

the venue. This is important as crowds usually exist in certain areas or at particular times of the day. Places where crowd density rises up over time are likely to congest and need careful observations to ensure the crowd safety. Basically, crowd density surveillance and estimation can be a good solution for management and controlling the crowds safety.

The results of the estimations obtained during the tests allow us to consider both methods successfully. While the statistical method reached quite good estimation rates (around 92 percent) for most groups, the spectral method illustrated small deviations between the best and the worst estimations, reaching on average almost the same rates of correct estimation obtained by the statistical method.

5. Method of Getting the Pedestrian Speed

As speeds are hard to observe, walking times are measured, from which walking speeds are derived. In addition to walking times and pedestrian densities, other variables need to be considered to complete the input of the simulation model (such as the number of in and out going pilgrims and the configuration of the structure during the rush hour at the Hajj period). The observables are the walking time, velocities and the corresponding densities of the pilgrims performing their Tawaf and Sa'y. The movements of the pilgrims going in and out of the Haram give us data to calculate the flux related to the Tawaf. The distribution of both in and out going pilgrims over the Haram can be derived from this data. The second type of observation concerns individual walking times. In order to measure the pilgrims' walking times in and out of the Haram, pilgrims are recorded from the moment they start walking from one spot to another, either on the piazza or going up the stairs. The start and duration of activities, such as Tawaf or Sa'y, are measured too. Finally, locations of origin, destination and possible activities of the pilgrims are registered. To do this, the piazza is divided into small areas with a length of 5×5 meters. We also record the movements of the pilgrims at specific moments, such as prayer times when the number of pedestrians increases dramatically. Therefore, cumulative flow curves can be constructed, out of which densities can be derived. These curves can be compared with the reference curves of Predtetschenski-Milinski (Predtetschenski & Milinski, Russ: 1969, Germ: 1971).

5.1. Subject Selection

Data was collected on a specific subject group of pedestrians who appeared to be 40 years old of age or older. On the roof of the Mataf area we selected our tracking subjects, consisting of adult men, women and people in wheelchairs. The following individuals were specifically not considered:

- Children under 13 years of age,
- Pedestrians carrying children, heavy bags, or suitcases,
- Pedestrians holding hands or assisting others across the Mataf,
- Pedestrians using a quad pod cane, walker, two canes, or crutches.

To accurately quantify the normal walking speeds of the various subject groups, pedestrians who exhibited any of the following behaviour were also not considered:

- Crossing of the Mataf path diagonally,
- Stopping or resting in the Mataf area,
- Entering the roadway running (anything faster than a fast walk),

The pedestrian sex (male or female) of each individual in the Mataf area was recorded, as well as whether he or she was walking alone or in a group. The group size was also noted when applicable. A group was defined by two or more pilgrims walking the Mataf trajectory at about the same time, regardless of whether or not they were apparently friends or associates. In the Mataf area, the pedestrian groups can reach 30 pilgrims walking together in the pedestrian stream. In addition, subjects paths were monitored to determine when they started and ended their Tawaf. Being inside the Mataf was defined as being within or on the painted Tawaf walking lines. Other pedestrian behaviour was recorded when it occurred:

- Confusion (hesitation, sudden change in direction of travel or change of point of interest) exhibited before walking,
- Confusion exhibited after entering the Mataf trajectory,
- Cane use,
- Following the lead of other pedestrians,
- Stopping in the walking path during the Tawaf movement,
- Difficulty going into Mataf,
- Difficulty going out of the Mataf.

Several methods were developed to check the accuracy and performance of walking speed estimation abilities of the observers. First, the walking speed was measured at the same time by three observers, then correlations between the estimates of all observers were determined. In particular, the walking velocity of one pilgrim was measured by the three observers and the mean value was taken. The results of these verification procedures are discussed after the next section.

5.2. Manual Methods

From our video recordings we chose places between two minarets as references (see [Figure 11](#)). As the dimensions of the mosque were known, we then established a grid of regular cells covering all of the Mataf area, each one having a size of 5 m × 5 m (see [Figure 12](#)). The distance between the two minarets was known. Pedestrian crossing times were measured with a digital timer and an electronic stopwatch was implemented and synchronized with the timer of the video recorder. The watch was started as the subject stepped off the first minaret and stopped when the subject stepped out on the opposite minaret after crossing all the distance between the two minarets.

5.2.1. Verification of Observer Walk-Speed Estimates and Start-Up Time Measurement

From the roof of the Mosque every pedestrian can be identified. To establish the ability of the field observers to identify the fitness level or the age of pedestrians with high accuracy a simple verification procedure was performed. The age estimation and the level of fitness of the pedestrians was based on their walking speed. It is a physio-medical fact that older pedestrians walk more slowly than younger ones (this is easily supported by field data), however, the published or already existing data on walking speeds and start-up times (i.e. the time from the beginning of a Tawaf movement until the pedestrian steps off the Mataf) have many shortcomings. Here we consider the complicated movement of the Tawaf and the human error rate of the observer. The walking speed on the Mataf area can be affected by many factors, one of the relevant factors is the age of the pedestrian. This demonstrates that the observations are quite good at identifying older pedestrians or pedestrians with fitness deficiency or physical health problems. A digital stopwatch was integrated with the video recording sophisticated for the measurements of pedestrian crossing times. The crossing times of the same pilgrims were measured during five rounds of the Tawaf and the average value was determined.

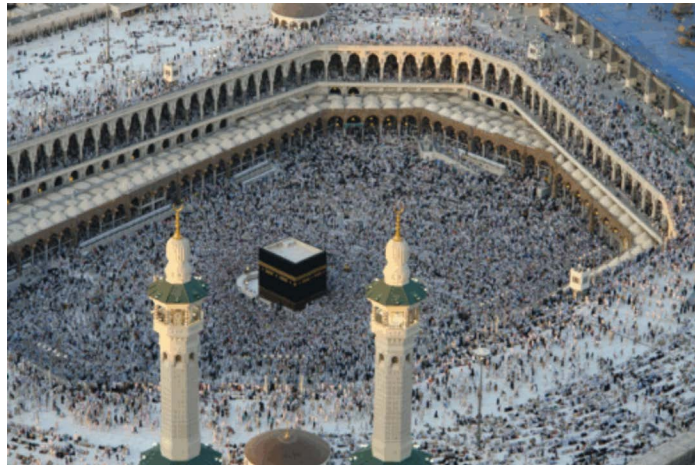


Figure 11. Overview of the piazza of the Haram, the place where our observations are made. With a digital clock the individual walking times t_p are measured. Since the distance between the two minarets is known from the architectural plan of the Haram, the average of the local pedestrian velocities $v(\mathbf{r}, t) = \|\mathbf{v}(\mathbf{r}, t)\|$ can be determined.

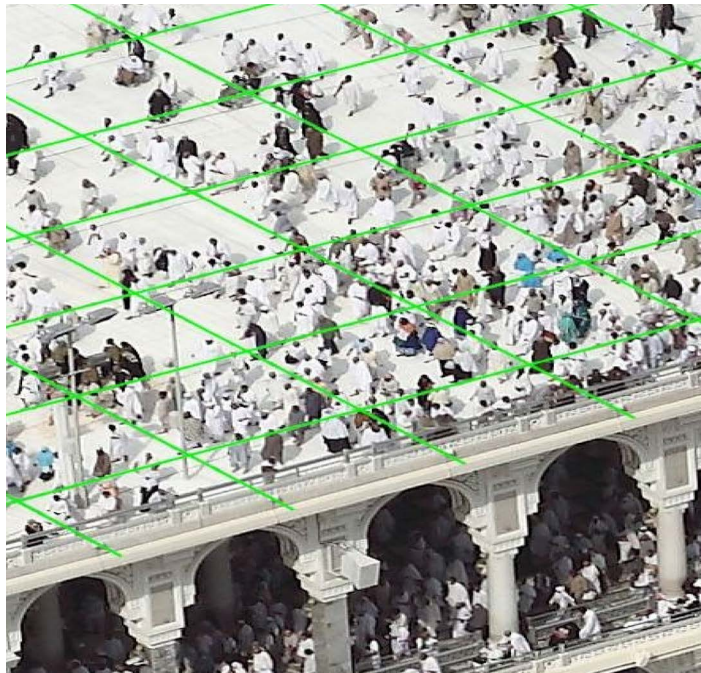


Figure 12. Grid of regular cells with dimension of $5 \text{ m} \times 5 \text{ m}$. With the help of the regular cells and the distance between two minarets in the Haram (see **Figure 11**) the (individual) walking times t_p are determined and the average of the local speeds $v(\mathbf{r}, t) = \|\mathbf{v}(\mathbf{r}, t)\|$ is calculated. The average walking speed for male pedestrians is 1.37 m/s , female 1.22 m/s and for people moving on wheelchairs 1.534 m/s .

5.2.2. Pedestrian Walking Speeds Results

This research also examined the impact of the building layout on the pedestrian speed distribution and the pedestrian density of pilgrims performing the Tawaf movement around the Kaaba. The set of data of pedestrian walking speeds which are obtained through analysing video recording using a set of statistical techniques are

displayed in **Figures 13(a)-(c)**. The results reveal that walking speed seems to be following a normal distribution no matter of male, female, older or younger. The average speed of young people is dramatically larger than that of older people, and the average speed of male is slightly larger than that of female. The width of the obtained curves is related to the different standard deviations.

The mean computed walking speed represented the speed that 85 percent of pedestrians did exceed. A total of 250 pedestrians were observed, including 100 male pedestrians of about 60 years of age, 100 women pedestrians and 50 wheelchair pedestrians. This data described all of the pedestrians observed: those walking in the center of the stream and those walking by the edge of the Mataf trajectory. As was subsequently described, those who were walking by the edge of the Mataf tended to walk more quickly. All observed pedestrians moved in a rotational motion around the Kaaba counter-clockwise (Tawaf), in compliance with the pilgrim stream.

The mean walking speed for male pedestrians was 1.37 m/s and 1.22 m/s for female pedestrians. In conjunction with pilgrims' age, the mean walking speed for younger pedestrians was 1.48 m/s and 1.20 m/s for older male pedestrians. The results revealed that the average walking speed for young women were 1.32 m/s and 1.12 m/s for old women. This means:

- Young male pedestrians had the fastest mean walking speed [1.48 m/s] and older females had the slowest [1.12 m/s]. The differences were between young men and young women [0.16 m/s] and between older men and older women [0.1 m/s], this result showed a little deviation that can be traced back to the fitness level of pedestrian or other factors, in the normal condition were approximately the same. The mean walking speed for the younger pedestrians ranged from 1.37 to 1.57 m/s across all conditions, with an overall mean speed of 1.48 m/s. The means for the older pedestrians ranged from 0.97 m/s to 1.26 m/s, with an overall mean speed of 1.18 m/s. For design purposes a mean speed of 1.33 m/s appeared appropriate;
- Locations by the edge of the Mataf had faster walking speeds because such locations have a lower pedestrian density. It was clear that the pedestrians near the Kaaba had a short walk path but in this places densities of 7 to 8 persons/ m² can be exceeded, making the movement of pilgrims very slow and turbulent;

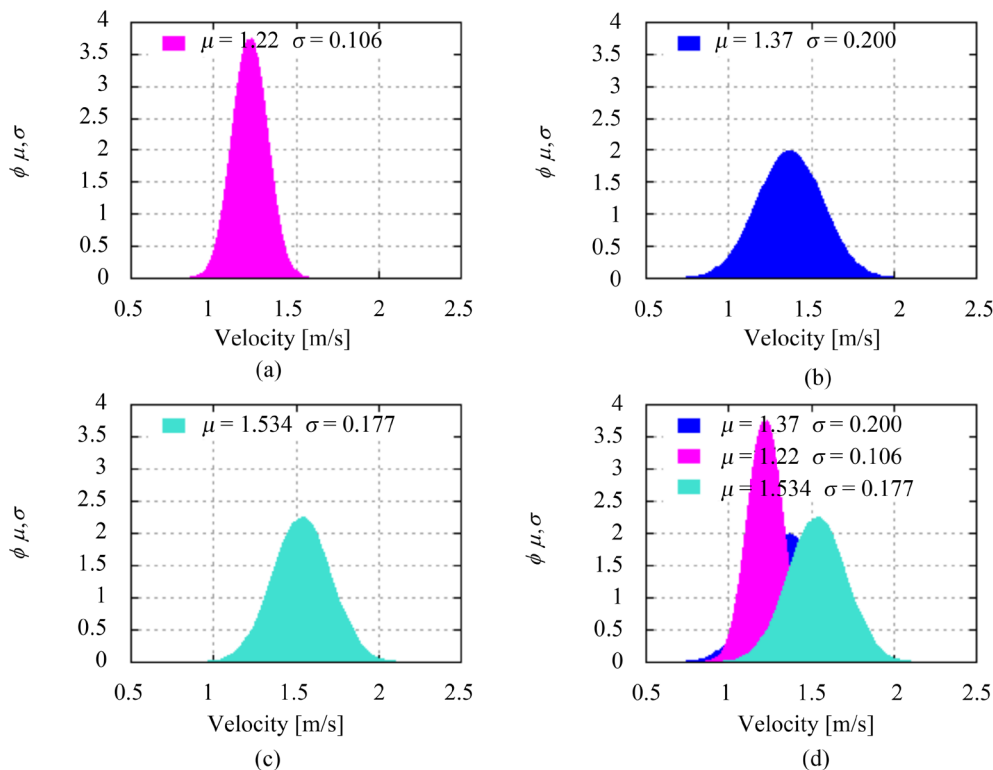


Figure 13. Walking speed distribution in Tawaf movement. (a) women, (b) men, (c) wheelchairs, and (d) shows a comparison of the three distributions. The average walking speed for female pedestrians is $m = 1.22$ m/s with a standard deviation of $s = 0.106$ while for male pedestrians $m = 1.37$ m/s and $s = 0.200$ and for pilgrims on wheelchairs $m = 1.534$ m/s and $s = 0.177$.

- Places situated further away from the Kaaba wall also tended to be associated with faster walking speeds. It was known from other fundamental diagrams that pedestrians tended to walk faster along a free walkway. As might be expected the walking speeds associated with various factors. The motion of a single individual at any given time and the direction and speed resulted in a long list of possible (and very likely conflicting) forces and circumstances.

The data taken show that each of the locations and surrounding factors have a significant effect on the behaviour and walking speed of the pilgrims on the Mataf area, not forgetting that the age of the pedestrians play a significant role in the Tawaf movement and density peaks and jams are caused by pilgrims of age 70 and more. For approximately one half of the location, the factors examined there also showed an important correlation between pedestrian age, the location and the mean walking speed of the pilgrims. This finding is consistent with results published by Knoblauch (Knoblauch et al., 1996).

The walking speed of pilgrims shows statistically significant variations across a variety of sites, times and environmental conditions (pedestrian density on the Mataf area). On the roof of the Mosque the pilgrim density is low and every pedestrian can walk with his desired velocity. However, the mean walking speed data is explicit by clustered for both pedestrians sex, men and women, independent of the age of the pilgrims are considered.

5.3. Automatic Estimation of Pedestrian Walking Speeds

There exist numerous methods that track the movement of single individuals by inspecting their orientation and limb positions.

This section highlights a real-time system for pedestrian tracking from sequences of high resolution images acquired by a stationary (high definition) camera. The objective was to estimate pedestrian velocities as a function of the local density. With this system the spatio-temporal coordinates of each pedestrian during the Tawaf ritual were established. Processing was done through the following steps:

- Existing footage was loaded onto a 3D program as a backplate,
- From several provided 2D- architectural drawings, a 3D model of the mosque was built,
- A virtual camera was matched in position, rotation and focal length to the original camera so that the features of the 3D-model matched the features positioned on the filmed mosque,
- Individual features were identified by eye, contrast is the criterion,
- We do know that the pilgrims walk on a plane, and after matching the camera we also obtained the height of the plane in 3D-space from our 3D model,
- A point object was placed at the position of a selected pedestrian. During the animation we set multiple animation-keys (approx every 25 to 50 frames (equals 1 to 2 seconds)) for the position, so that the position of the point and the pedestrian overlay nearly all the time,
- By evolving the point with time we obtained the distance travelled, by measuring the distance from frame to frame. We also knew the time elapsed from the speed per frame, and hence the speed could be calculated.

5.4. Analysis

From **Figure 14** and **Figure 15** we see that the edge of the Mataf moves faster than the center, this phenomenon being known as the Edge Effect. The Edge Effect occurs when the edges of a crowd move faster than the center of the crowd. The density becomes higher and higher as one moves from the edge of the Mataf towards the center. This phenomenon is explained by the fact that all pilgrims want to be near the Kaaba wall. As a result, we find the density near the Kaaba to be the maximum density. This data can be used in validating of simulation tools. The mean walking speed for a group of pedestrians moving in the pilgrim stream around the Kaaba was 1.0816 m/s at the edge of the Mataf and it was 0.3267 m/s for the same pedestrians groups moving inside the Mataf. These findings agree well with the statistical results discussed in a previous section.

6. Comparison of Walking Speeds

One of the must-have results is to compare the mean values and variances of walking speeds in both observations and simulation results. A distinction will be made for walking speeds inside and outside of the Mataf platforms. We made a comparison between our plots derived from the video observation and the fundamental

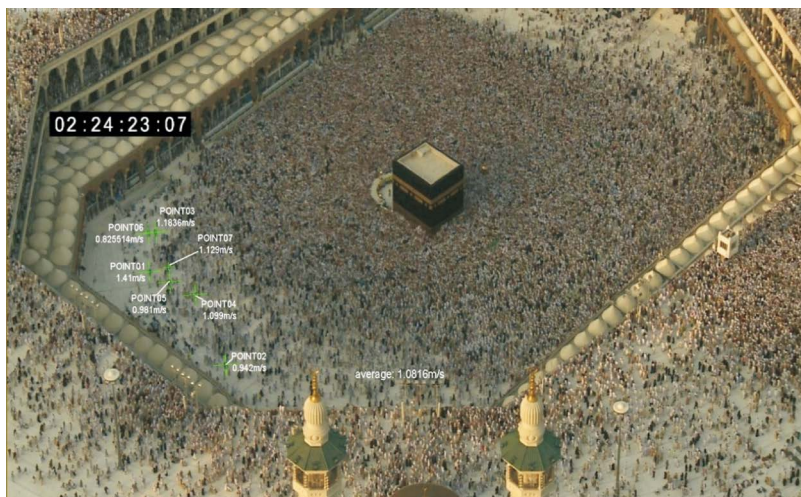


Figure 14. Pilgrims’ walking speeds on the edge of the Mataf. The average walking speed is 1.0816 m/s.

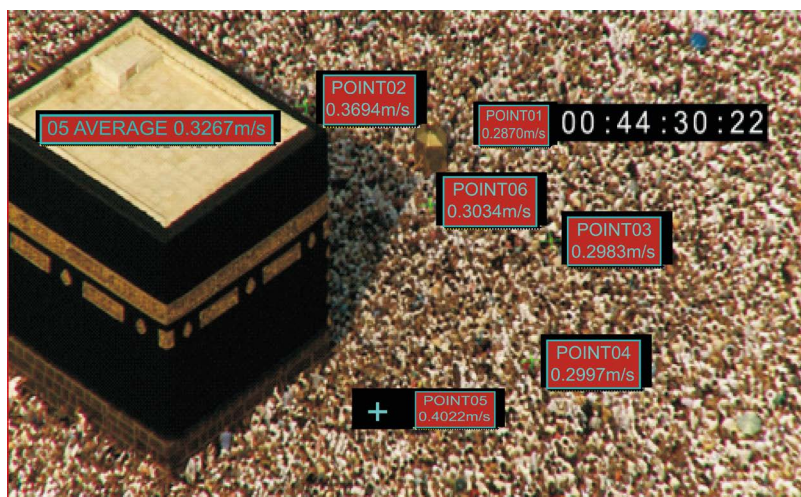


Figure 15. Pilgrims’ walking speeds in the Mataf area (near the Kaaba wall). The average walking speed is 0.3267 m/s.

diagrams of (cf. **Figure 16**):

- Walking speeds:
 - On the edge of the Mataf (free flow speed) where the pedestrian density is lower than 3 persons/m²,
 - On the center of the Mataf,
 - On the Mataf inside near the Kaaba wall where the pedestrian density attains extreme levels (8 - 9 persons/m²).

All well-known fundamental diagrams predict the same behaviour and have the same properties: speed decreases with increasing density. So the discussion above indicates there are many possible reasons and causes for the speed reduction. For example there is a linear relationship between speed and the inverse of the density for pedestrians moving in a straight way (Seyfried et al., 2005). However the pedestrian walking speed can be affected by internal and external factors (such as the amount of pedestrian inflow and outflow as well as the configuration of the infrastructure) not to forget the physiology of the human body. It is found that individuals walk faster in outdoor facilities than in corridors (Lam & Cheung, 2000). According to Predtechenskii and Milinskii (PM) the average walking speed depends on the the walking facility (Predtetschenski & Milinski, Russ: 1969, Germ, 1971). In other circumstances Weidmann confirmed a linear relationship between the step size length of walking pedestrians and the inverse of the density (Weidmann, 1993). The small step size means low pedestrian

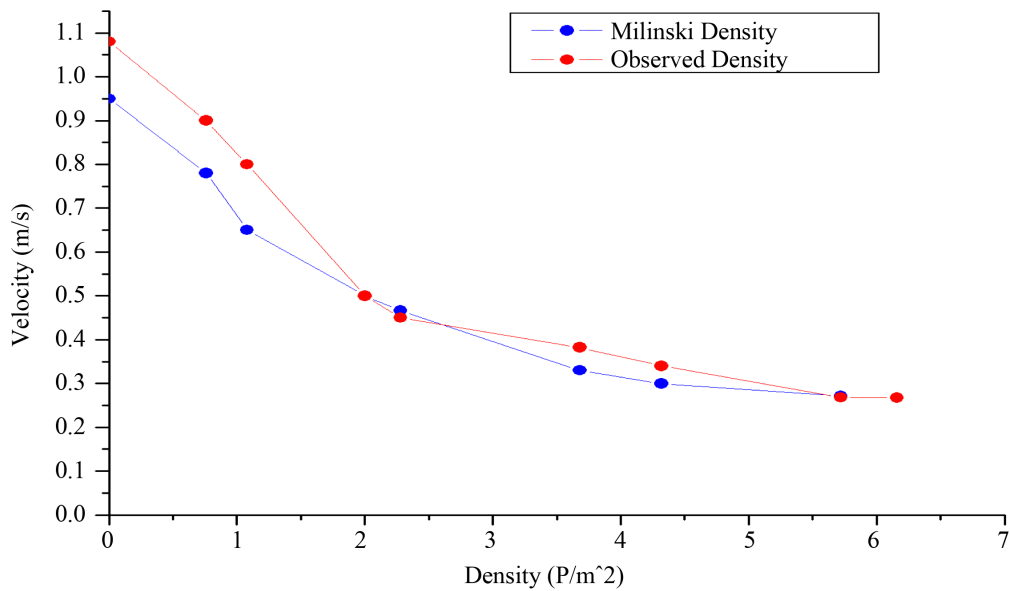


Figure 16. The Velocity-Density-Diagrams half-hour after Mid-Day Prayer ($t = t_{\text{half-hour after Mid-Day}}$). Average of the local speeds $v(r,t)$ as a function of the local density $\rho(r,t)$. Our own data are shown as red points. The blue points correspond to the data obtained by (PM) (Predtetschenski & Milinski, Russ: 1969, Germ: 1971). The difference in velocity at lower densities can be explained by the fitness level of pedestrians.

velocity, caused by reduction of the available space with increasing density. The discussion above shows that there are many possible factors influencing the fundamental diagram. To identify these factors, it is necessary to exclude as many influences of measurement methodology and short range fluctuations from the data. **Figure 16** shows the average local speed $v(r,t)$ as a function of the local density $\rho(r,t)$ half-hour after Mid-Day Prayer ($t = t_{\text{half-hour after Mid-Day}}$). Our own data is shown as red points. The blue points correspond to the Milinski fundamental diagram. Moreover investigation data analysing the Mataf area represented by blue points in figure 16 and showed that a reduction of the available navigation space illustrates the causes responsible for the speed reduction with density in pedestrian movement. The small deviation in pedestrian walking speed at lower density can be explained by the fitness level of the pedestrian.

7. Movement Recognition

In the literature, there is a large number of approaches on detection and tracking of moving objects from video images. Spatio-temporal analysis has, in the past, been used to recognize walking persons, where subspaces in the video are treated as spatio-temporal volumes (Ricquebourg & Bouthemy, 2000). Application of a Fourier transform to this data can then identify data relating to movement across the volume. This approach allowed pedestrian trajectories to be reconstructed from video with high precision, taking advantage from the methods and the high developed computational technology. The common approach to detect movement is to produce comparison images (an image representing the different details between two images) since this is computationally efficient (Masoud & Papanikolopoulos, 2001). These comparison images can then be computed further to estimate movement vectors that describe the motion of drop-shaped objects captured in the respective images. Murakami and Wada demonstrate another method, filing the difference frame, and instead compare the properties of drops identified in consecutive frames (Murakami & Wada, 2000). A drop that is close to the position of a drop in a previous frame, and shares similar dimensions, is likely to refer to the same figure. Motion vectors are also used to find drop segmentation, which are subsequently merged or separated for the purpose of analysis. The same approach is applied to a 2D image to determine movement in 3D space. Extrapolating the movement of pedestrians in 3D space from a 2D image allows for a far greater understanding of the interactions between entities, but does require exceptional calibrations of equipment for complete accuracy.

The Murakami and Wada approach can be used to analyse low-quality video streams due to the frame-differencing algorithm and some trigonometry. Determining 3D motion does require precise knowledge of the angle and position of the camera, in addition to the basic topology of the scene being analysed. But 2D paths are easy to identify without these details, (see [Figure 17](#)).

In [Figure 17](#) we show the path of individuals within the crowd. One clearly recognizes that the movement around the Kaaba is not a circle movement. The tracking of a single individual in the pilgrim stream indicates some oscillation movement around the main path of the individual. It is caused by the physical repulsive and attractive forces acting on the individual. Physical forces become important when an individual comes into physical contact with another individual/obstacle. When a local density of 6 persons per square meter is exceeded, free movement is impeded and local flow decreases, causing the outflow to drop significantly below the inflow. This causes a higher and higher compression in the crowd, until the local densities become critical in specific places on the Mataf platform.

8. Analysis of the Pilgrims Movement on the Mataf

In the Mataf everything is dense and we have a compact state. The pilgrims have body contact in all directions and no influence on their movement; they float in the stream. This forms structures and turbulences in the flow. These turbulences can be well observed in our video recording. Density and velocity can also be seen. These observed Hajj rituals, especially the Mataf, showed some critical points in the motion of the pilgrims that we had not paid much attention to before. For example: the edge effect, density effect, shock-wave effect, etc. and phenomena like these influence the restraint of the motion and are very important to be considered.

Our video analysis shows that the pedestrian density decreases with the distance from the Kaaba wall (cf. [Figures 6-8](#)). It is the same as the real behaviour of pilgrims on the Mataf ritual (all pilgrims want to be near to the Kaaba wall). Our video analysis about the Mataf area indicates that, even at extreme densities, the average local speeds and flows stay limited. This extremely high local density causes forward and backward moving shock-waves, which could be clearly observed in our video. We can see a kind of oscillation on the pilgrims paths around the Kaaba, this oscillation is caused by shock-waves and is affected by the repulsive forces between the pedestrians in high density crowds (see [Figure 17](#)).

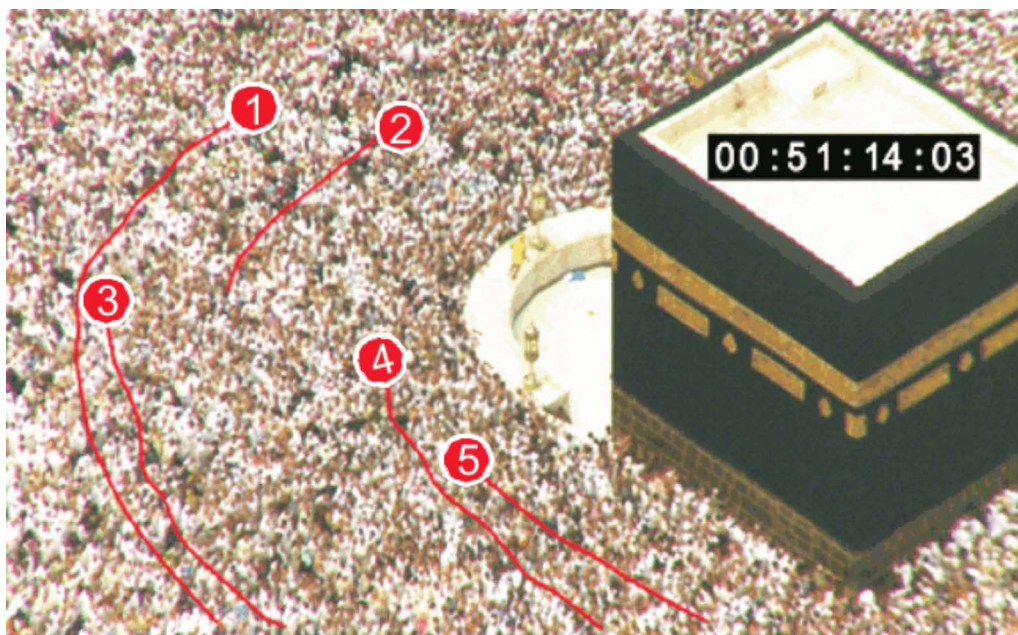


Figure 17. Pilgrims paths. With a new computer algorithm developed during this research, the trajectories or movements of pedestrians across the infrastructure over time are determined. Microscopic pedestrian fields require large amounts of trajectory data of individual pedestrians. Every red solid curve corresponds to one pedestrian trajectory. The oscillation in the pilgrims paths results from the huge pedestrian forces acting on every individual in the crowd.

9. Conclusions and Possible Improvement

One of the significant challenges in the planning, design and management of public facilities subject to high density crowd dynamics and pedestrian traffic are the shortcomings in the empirical data. The collected data concerning crowd behaviour using different techniques (image processing) and analysis of ordered image sequences obtained from video recording is increasingly desirable in the design of facilities and long-term site management. We have investigated the efficiency of a number of techniques developed for crowd density estimation, movement estimation, critical places and events detection using image processing. In the above sections and within this investigation we have presented techniques for background generation and calibration to improve the previously developed simulation model.

Even though extracting information about human characteristics from video recording may still be in its infancy, it is important to mention that the field of human motion analysis is large and has a history traced back to the work of Hoffman and Flinchbaugh (Hoffman & Flinchbaugh, 1982). In the field of pedestrian detection techniques, moreover in the big area of computer vision, many problems have accumulated. In the human motion analysis, and also in the problem of the detection of moving objects, there still remains other problems, namely to recognize, categorize, or analyse the long-term pattern of motion. The inspection of the literature in the last decade indicates increasing interest in event detection, video tracking, object recognition, because of the clear application of these technologies to problems in surveillance. Recently many methods have been developed to extract information about moving object like speed and density. Almost all these systems require complex inter- mediate processes, such as reference points on the tracked objects or the image segmentation. One limitation of this current system is that the detection failures for these intermediates will lead to failure for the entire system.

Improvement of an algorithm to be able to reproduce traffic flow and to help in the microscopic pedestrian data collection is very essential. Moreover the automatic video data collection will highly enhance the achievement of a system for higher pedestrian traffic densities.

Acknowledgements

I would like to express my sincerest thanks and gratitude to Prof. Dr. G. Wunner for a critical reading of the manuscript, for his important comments and suggestions to improve the manuscript. Many thanks to Dr. H. Cartarius for his support during writing this work.

References

- Boltes, M., Seyfried, A., Steffen, B., & Schadschneider, A. (2010). Automatic Extraction of Pedestrian Trajectories from Video Recordings. In W. W. F. Klingsch, C. Rogsch, A. Schadschneider, & M. Schreckenberg (Eds.), *Pedestrian and Evacuation Dynamics 2008* (pp. 43-54). Heidelberg: Springer Berlin Heidelberg. http://dx.doi.org/10.1007/978-3-642-04504-2_3
- Bulpitt, A. J., & Sumpster, N. (1998). Learning Spatio-Temporal Patterns for Predicting Object Behaviour. *Proceedings of the 9th British Machine Vision Conference (BMVC)*, Southampton, 649-658.
- Chattaraj, U., Seyfried, A., & Chakroborty, P. (2009). Comparison of Pedestrian Fundamental Diagram across Cultures. *Advances in Complex Systems*, 12, 393. <http://dx.doi.org/10.1142/S0219525909002209>
- Fruin, J. J. (1987). *Pedestrian Planning and Design*. Mobile, AL: Elevator World, Inc.
- Fruin, J. J., American Society of Mechanical Engineers, & American Society of Mechanical Engineers, Standing Committee on Transportation (1970). *Designing for Pedestrians: A Level of Service Concept*. Ann Arbor, MI: University Microfilms, Inc. http://books.google.de/books?id=LM_EtgAACAAJ
- Heisele, B., & Woehler, C. (1998). Motion-Based Recognition of Pedestrians. *Proceedings of 14th International Conference on Pattern Recognition*, Brisbane, 16-20 August 1998, 1325-1330. <http://dx.doi.org/10.1109/ICPR.1998.711946>
- Helbing, D., Al-Abideen, H. Z., & Johansson, A. (2007). The Dynamics of Crowd Disasters: An Empirical Study. *Physical Review E*, 75, Article ID: 046109. <http://dx.doi.org/10.1103/PhysRevE.75.046109>
- Hoffman, D. D., & Flinchbaugh, B. E. (1982). The Interpretation of Biological Motion. *Biological Cybernetics*, 42, 195-204.
- Hoogendoorn, S. P., & Daamen, W. (2005). Pedestrian Behavior at Bottlenecks. *Transportation Science*, 39, 147-159. <http://dx.doi.org/10.1287/trsc.1040.0102>

- Johansson, A., Al-Abideen, H. Z., Al-Bosta, S., & Helbing, D. (2008). From Crowd Dynamics to Crowd Safety: A Video-Based Analysis. *Advances in Complex Systems*, 11, 497-527. <http://dx.doi.org/10.1142/S0219525908001854>
- Knoblauch, R. L., Pietrucha, M. T., & Nitzburg, M. (1996). *Field Studies of Pedestrian Walking Speed and Start-Up Time* (pp. 27-38). Transportation Research Record 1538, Washington DC: National Research Council, Transportation Research Board.
- Lam, W. H. K., & Cheung, C. Y. (2000). Pedestrian Speed/Flow Relationships for Walking Facilities in Hong Kong. *Journal of Transportation Engineering, ASCE*, 126, 343-349.
- Li, X., Shen, L., & Li, H. (2006). Estimation of Crowd Density Based on Wavelet and Support Vector Machine. *Transactions of the Institute of Measurement and Control (London)*, 28, 299-308.
- Marana, A. N., Costa, L. F., Lotufo, R. A., & Velastin, S. A. (1998). On the Efficacy of Texture Analysis for Crowd Monitoring. *Proceedings of SIBGRAP'98. International Symposium on Computer Graphics, Image Processing, and Vision*, Rio de Janeiro, 20-23 October 1998, 354-361.
- Masoud, O., & Papanikolopoulos, N. P. (2001). A Novel Method for Tracking and Counting Pedestrians in Real-Time Using a Single Camera. *IEEE Transactions on Vehicular Technology*, 50, 1267-1278.
- Murakami, S., & Wada, A. (2000). An Automatic Extraction and Display Method of Walking Persons' Trajectories. *Proceedings of 15th International Conference on Pattern Recognition*, 4, 611-614.
- O'Flaherty, C. (1996). *Transport Planning and Traffic Engineering*. Engineering Village. Oxford: Taylor & Francis. <http://books.google.de/books?id=BjJPAAAAMAAJ>
- Papageorgiou, C., & Poggio, T. (1999). Trainable Pedestrian Detection. *Proceedings of 1999 International Conference on Image Processing*, 4, 35-39.
- Predtetschenski, W. M., & Milinski, A. I. (Russ: 1969, Germ: 1971). *Personenströme in Gebäuden*. Berlin: Staatsverlag der Deutschen Demokratischen Republik.
- Ricquebourg, Y., & Bouthemy, P. (2000). Real-Time Human Figure Control Using Tracked Blobs. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22, 797-808.
- Sen, G., Wei, L., & Ping, Y. H. (2009). Counting People in Crowd Open Scene Based on Grey Level Dependence Matrix. *IEEE International Conference on Information and Automation*, Zhuhai, 22-24 June 2009, 228-231.
- Seyfried, A., Steffen, B., Klingsch, W., & Boltes, M. (2005). The Fundamental Diagram of Pedestrian Movement Revisited. *Journal of Statistical Mechanics*, 2005, P10002. <http://dx.doi.org/10.1088/1742-5468/2005/10/P10002>
- Tsuchikawa, M., Sato, A., Koike, H., & Tomono, A. (1995). A Moving-Object Extraction Method Robust against Illumination Level Changes for a Pedestrian Counting System. *Proceedings of International Symposium on Computer Vision*, 563-568.
- Verona, V., & Marana, A.N. (2001). Wavelet Packet Analysis for Crowd Density Estimation. *Proceedings of the Listed International Symposium on Applied Information*, Innsbruck: Aeta Press, 535-540.
- Weidmann, U. (1993). *Transporttechnik der fussgänger*. Technical Report Schriftenreihe des IVT Nr. 90, Zürich: Institut für Verkehrsplanung, Transporttechnik, Strassen-und Eisenbahnbau, ETH Zürich, Zweite, ergänzte Auflage.
- Wu, X. (2006). Crowd Density Estimation Using Texture Analysis and Learning. *IEEE International Conference on Robotics and Biomimetics*, Kunming, 17-20 December 2006, 214-219.
- Zhang, Z., & Li, M. (2012). Crowd Density Estimation Based on Statistical Analysis of Local Intra-Crowd Motions for Public Area Surveillance. *Optical Engineering*, 51, Article ID: 047204. <http://dx.doi.org/10.1117/1.OE.51.4.047204>