Extracting Microscopic Pedestrian Characteristics from Video Data

Results from experimental research into pedestrian walking behavior

Serge P. Hoogendoorn[1] (s.hoogendoorn@ct.tudelft.nl)
Winnie Daamen (wdaamen@hr.nl)
Piet H.L. Bovy (bovy@ct.tudelft.nl)

Transportation and Traffic Engineering Section
Faculty of Civil Engineering and Geosciences
Delft University of Technology

Abstract. Insight into walking behavior is essential for theory and model development describing the behavior of individual pedestrians. In turn, microscopic pedestrian simulation models can be used to test and compare different infrastructure designs, both from the perspective of efficiency and safety. To calibrate these microscopic models, detailed data is required. However, currently such data is not available.

This paper discusses an approach to extract individual pedestrian data from digital video footage. In this specific case, the footage was collected during a large scale walking experiment performed at the Delft University of Technology. In this experiment, a group of approximately 80 pedestrians conducted different tests, while being observed using a video camera mounted at the ceiling of the hall.

In the first phase of the pedestrian tracking process, different image processing operations, such as radiometric correction and correction for lens distortion are applied. Next, the pedestrians are detected by identifying the special colored caps the pedestrians were wearing during the experiment. This detection occurs in a special zone where conditions are such that pedestrian detection is optimal. Next, pedestrian tracking is achieved by application of a newly developed tracking technique. This is achieved by minimizing the so-called merit-function. A Kalman filter is applied to reduce the errors made during tracking.

The paper concludes by showing some of the tracking results, and accuracy estimates. It turns out that the pedestrian trajectories can be determined with high accuracy. Also, some interesting findings pertaining to the behavior of pedestrians are reported.

Keywords: image analysis, walking behavior, micro-simulation models, pedestrian detection, pedestrian tracking.

Word count	
Abstract	241
Main text	5378
Figures and tables (7 x 250)	1750
Total	**7369**

BACKGROUND

To optimize the design of airports, public transit stations, stadiums, shopping malls, etc., theory and models predicting spatial-temporal pedestrian flow patterns and individually experienced walking conditions, in terms of walking times, levels-of-service, etc., are important. This holds equally for timetable design, where walking times experienced by public transit riders when accessing or transferring play an important role. Analysis tools can support infrastructure designers as well as public transport planners to optimize their design, making it more efficient, more comfortable, and safer. This is why different researchers have developed a dedicated comprehensive dynamic theory for pedestrians with related models.

[1] Corresponding author.

Some of the models proposed recently describe and predict the behavior of the individual pedestrians, that is, they are microscopic. These models are usually established from different behavioral hypotheses. For instance, Helbing et al. (*1*) assume that pedestrians can be described by particles on which so-called *social-forces* act. These social forces determine the walking behavior and interaction between pedestrians and obstacles under varying circumstances. Hoogendoorn and Bovy (*2*) assume that pedestrians can be compared to autonomous controllers, optimizing the predicted cost of their walking actions. Using a Cellular Automata model, Blue and Adler (*3*) use simple production rules to describe pedestrian behavior under varying circumstances.

The calibration of these microscopic simulation models requires a high level of detail in data collection in order to be useful. Brackstone and MacDonald (*4*) convincingly argue that for the calibration of 'true' microscopic simulation models – in this case, of vehicular traffic – detailed information on the interaction between vehicles is required: to improve modeling of car-following and lane-changing behavior, microscopic behavior of vehicles *compared to vehicles* upon which they react and anticipate, is needed. The fundamental character of this issue expresses itself in the *high level of detail required* in the research data. We argue that this holds equally for microscopic pedestrian models, and thus that this type of fundamental research is not possible without data that consists of complete pedestrians trajectories.

Not only the microscopic behavior of pedestrians is important. In order to gain insight into macroscopic flow characteristics, phenomena such as dynamic lane formation, of the characteristics of crossing pedestrian flows, capacity of bottlenecks, detailed information of pedestrian flows is important as well.

This is why the Transportation and Traffic Engineering Section of Delft University of Technology has been performing unique experimental research aimed at gaining more insight into microscopic pedestrian behavior and the characteristics of the pedestrian flows. During the research, a large group of pedestrians divided into smaller controllable groups, was instructed to perform ten different walking experiments. The experiments were conducted in a large hallway of the University building. The motion of the individual pedestrians was observed using a digital video camera mounted on the ceiling of the building. The location of the camera allowed the researches to observe the pedestrians from above, enabling precise reconstruction of the pedestrian trajectories.

This contribution describes the approach that was developed to reconstruct the pedestrian trajectories from the digital video images, as well as the resulting microscopic pedestrian characteristics. To this end, first briefly the measurement set-up will be described. Secondly, an overview of the pedestrian tracking approach is given. In the subsequent sections, the main parts of the approach are described in detail. In the final section of the paper, the obtained microscopic data is presented.

STATE-OF-THE-ART REVIEW

Before describing the measurement and experimental set-up, we will briefly consider a brief (and incomplete) state-of-the-art in empirical and experimental pedestrian research. The second part of this section discusses some relevant publications in image processing and object detection and tracking.

Empirical and experimental pedestrian flow research

Available empirical and experimental knowledge on pedestrian walking behavior, and especially concerning the walking behavior of individuals, is increasing slowly. This section considers some research findings pertaining to the fundamentals of the walking process, as well as some characteristics of the pedestrian flow, with the aim to provide a brief overview of some interesting research findings. For more details, we refer to (*5*).

Goffman (*8*) describes how the environment of the pedestrian (infrastructure and pedestrians) is *continuously observed* through a mostly subconscious process called *scanning* in order to side-step small obstructions on the flooring and handling encounters with other pedestrians; the author also describes how pedestrians react to one another by bi-lateral, subconscious communication. This is how pedestrians are generally aware of the future behavior of other pedestrians in their direct environment.

Other researchers also stressed the importance of co-operation between pedestrians: by experimentation (letting researchers walk in a crowd), Sobel and Lillith (*9*) show that pedestrians are reluctant to unilaterally withdraw from an encounter until the last moment, possibly even resulting in physical contact between the pedestrians. This "brushing" sends a signal to the 'offender' to co-operate. Naturally, the predictions are limited with respect to the time-horizon as well in a spatial sense. The latter is also due to the limited ob-

servation range. Not all encounters between pedestrians will result in a co-operative decision, as the amount of space granted by a pedestrian depends on the cultural, social, and demographic characteristics of the interacting pedestrians (*9*), (*10*). Willis *et al.* (*11*) found similar results for one-to-many interactions: individuals tend to move for groups.

Other researchers have focussed on the *characteristics of the pedestrian flow,* such as capacity, and the fundamental diagram. The most comprehensive survey so far is the work of Weidmann (*7*). Amongst the pedestrian flow characteristics discussed by Weidmann are

1. Free walking speeds and their dependence on personal and trip characteristics, environmental and weather conditions, etc.
2. Relation between speed and density (fundamental diagram).
3. Capacities of pedestrian facilities.

Weidmann also considers the effect of *flow composition* (with respect to direction) on average flow speeds and capacity. Lam and Cheung (*14*) compare different speed-flow relations for different types of pedestrian facilities. They conclude that pedestrians walk faster on outdoor facilities than on indoor walkways. In line with other researchers, they observe that average speeds depend on the use of the walking facility. For instance, pedestrians in commercial areas walk faster than in shopping malls. Comparing their results with results from other countries leads them to conclude that Asian pedestrians require less personal space than Western pedestrians. More recent work of Lam *et al.* (*13*) examines the relation between walking speed and pedestrian flow under various bi-directional flow conditions for in-door walkways in Hong-Kong.

Observation techniques for pedestrian flows

Compared to vehicular traffic, observing pedestrians in either real-life or experimental conditions is more involved. For one, neither inductive loops nor pneumatic tubes can be effectively used to collect pedestrian data. In special situations, automated detection of pedestrians is possible. However, in general researcher need to resort to optical measurement devices (infrared, video), radar, or GPS-like techniques. The use of photographs of video has the advantage that the observations are not restricted to local measurements. As a result, instantaneous variables such as densities and space-mean-speeds can be determined from traffic data directly. An example of the use of photography, Lam and Cheung (*14*) use time-lapse photography to collect the raw data material. The images are processed manually to determine the required data. Clearly, this approach is very labour-intensive.

An interesting observation technique is used by Hughes *et al*. (*15*). To study pedestrian flows in Mekka, they use ‘probes’ (instrumented pedestrians) that walk along with the pedestrian stream, allowing to determine the trajectories of a number of pedestrians and inferring the average speed of the pedestrian flow from that. Willis *et al*. (*16*) discuss a similar approach to track pedestrians from video data.

In this paper we propose an approach to *automatically detect and track pedestrians* from a sequence of high-quality video images. It goes without saying that an automated procedure has enormous advantages over the labour-intensive manual data collection. This holds especially for the pedestrian experiments described here, where many hours of high density pedestrian traffic operations have been monitored.

For successful application of automated detection and tracking of pedestrians, well-known and tested techniques from photogrammetry (e.g. lens correction for pin-cushion distortion, radiometric correction, etc.) are combined with dedicated techniques that make use of the special features of the available video data.

Object detection and tracking

The number of publications on detection and tracking of moving objects from video images is large. Some examples are found in (*17*)-(*21*). Most of the approaches use an approach similar to the approach described here. First, the raw images are corrected (correction for lens distortion, orhtorectification, and radiometric correction). Next, the objects are detected, frequently by first determining the empty scene, which is then subtracted from the images from which the objects are to be detected. Then, objects are detected and tracked through a sequence of images.

Application of these techniques to the problem at hand will not yield optimal results, due to its specific characteristics. For one, on the contrary to most of the objects that are tracked, the shape of the pedestrians changes as they walk due to the movement of the limbs. Secondly, when traffic conditions become congested, the detection of individual pedestrians becomes very complicated. As a result, dedicated techniques are to be developed for this specific problem.

MEASUREMENT AND EXPERIMENTAL SET-UP

This section briefly the experimental and measurement set-up. For a detailed description of the walking experiment, its aims, and its limitations, see Daamen and Hoogendoorn (*5*).

Experimental set-up

Experimental research entails interfering with natural processes in order to obtain more insight into the causal relations between the independent process variables (stimuli) and the observed phenomena (response). By performing experiments we can determine the causes and relations that determine the behavior of pedestrians. Apart from the methodological advantages, experiments allow observations of conditions that are not available or are very difficult to observe in normal conditions. The process variables are both the input and output variables that are deemed relevant. It is important that in an early phase of the research, a clear distinction has to be made between primary and secondary factors. In this case, the important primary factors have been determined from expert knowledge (*5*) and literature surveys (*5,7-14*). Table 1 shows the different factors that have been determined for the walking experiments, where we have distinguished between *experimental variables* and *context variables*: while the former were influenced during the experiments, the latter are specific for the different pedestrians. Note that some of the context variables are in a way response variables and are therefore a subject for study. In illustration, during one experiment, some pedestrians were asked to walk as if they were in a hurry, thus affecting the desired speeds.

Table 1 Process variables and observed phenomena for pedestrian experiments determined from expert knowledge (a brainstorm session) and literature survey. For details, see (*5*).

Stimuli		**Response**	
Experimental variables	**Context variables**	**Microscopic**	**Macroscopic**
▪ Desired walking direction (destinations) ▪ Aggressiveness (in terms of the extent in which the free speed is maintained) ▪ Bottlenecks	▪ Free or desired speed of the individual pedestrians[2]. ▪ Age of the pedestrian. ▪ Gender of the pedestrians. ▪ Grouping behavior of acquainted walkers	▪ Walking speeds ▪ Walking directions ▪ Distances between pedestrians ▪ Relative positions of pedestrians ▪ Passing behavior ▪ Group formation	▪ Density ▪ Space-mean-speed ▪ Intensity ▪ Time-mean-speed ▪ Desired speed distribution

In the end, pedestrian trajectories for distinct pedestrian groups are determined from the video pictures. These trajectories are the most elementary and most valuable unit of analysis in traffic flow research, which provide all information required for analysis of both the microscopic and macroscopic characteristics.

Overview of experiments

In total, 10 different walking experiments where performed. In each of these experiments, approximately 60-90 pedestrians were involved. The participants were divided into 8 groups, which were given separate instructions (walk slowly, walk fast, etc.). The groups themselves were heterogeneous, and consistent of men, and women of different ages. The group participants were indicated by the color of their caps (red or green). The red caps convey the ordinary behaving pedestrians, while the green caps were pedestrians that had to follow specific instructions (walk aggressively, walk slowly, etc.). The groups were not used to indicate the walking direction, as this could be determined from the video images straightaway. Table 2 shows an overview of the performed experiments.

[2] Pedestrians were asked to walk aggressively (or just the opposite) and choose the free speed that fits their level-of-aggressiveness. Thus, the free speed is a context variable and not a (direct) experimental variable.

Table 2 Overview of walking experiments.

Experiment type	Uni-directional flows	Bi-directional flows	Crossing flows	Bottleneck flows
Walking area dimension	10 m × 4 m	10 m × 4 m	8 m × 8 m	10 m × 4 m
Experiment description	1. Homogeneous group (normal walking conditions) 2. 40% aggressive pedestrians (increased desired speed); 60% normal pedestrians 3. 40% timid pedestrians (decreased desired speed); 60% normal	4. 50% from East to West; 50% from West to East 5. 90% from East to West; 10% from West to East	6. 50% from East to West; 50% from North to South 7. 90% from East to West; 10% from North to South 8. 25% from East to West; 25% from West to East; 25% from North to South; 25% from South to North	9. Uni-directional flow through wide bottleneck (2 meter width) 10. Uni-directional flow through narrow bottleneck (1 meter width)

For the uni-directional, two-directional, and bottleneck experiments, a walking area of 10 m × 4 m was used; for the crossing pedestrian flows a walking area of 8 m × 8 m was used (see figures 1a and b)

Figure 1 Examples of pedestrian experiments; a) narrow bottleneck experiment, and b) four-directional crossing flow experiment. Yellow circles (and demarcation on the floor) indicate the respective walking areas. The figure also shows the so-called pincushion distortion and the slight rotation of the camera (around 1%).

Measurement set-up

The walking experiments were conducted in a large hallway. The ambient conditions were favorable (reasonably constant light intensity, little shadows), although there was some influence of the sun shining through the windows. The digital camera was mounted at the ceiling of the hallway, at a height of 10 m, observing an area of approximately 14 m by 12 m. A wide lens was used enabling the camera to view the entire walking area. The distortions in the pedestrian trajectories caused by the slight pincushion effect and camera rotation were corrected for after the trajectories could be determined from the video data (step 4 in approach described in the following section).

The digital camera has a resolution of 720 × 576 pixels, and was attached to a digital video recorder. The quality of the collected video footage was very high.

ALGORITHMIC APPROACH TO PEDESTRIAN TRACKING

In broad terms, the different steps of the algorithm are the following:

1. *Converting digital video to image sequences and radiometric correction.*
2. *Identification of candidate pedestrians positions using fuzzy clustering*, with the aim of recognizing the pedestrians, and identify the different pedestrian groups, based on the colors of the caps. This pedestrian detection occurs in the so-called *detection zone*, where lighting conditions are optimal for recognition of the cap colors.
3. *Tracking detected pedestrians*, i.e. identifying subsequent pedestrian locations in the subsequent images by minimizing the merit function reflecting the matching between a subimage in image *i*-1 and image *i*.
4. *Mapping image coordinated to terrestrial coordinates*.
5. *Filtering the pedestrian trajectories.* Use the candidate pedestrian locations identified for the consecutive video images to establish the most likely trajectory, based on Kalman filtering techniques.
6. *Visualization of the trajectories.*

In the first step, the image sequences are converted from the digital video footage. Furthermore, these images are enhanced to improve the pedestrian detection. This is achieved by reducing temporal fluctuations in the image intensity (radiometric correction). To decrease computation time, it was decided that correction for rotation and the pincushion effect would be performed after the pedestrian tracking process was completed (i.e. during step 4).

Step 2 is more complicated, involving the identification of the individual pedestrians from the remainder of the scene. Several options have been considered here. However, the special properties of the study set-up (pedestrians with different tasks wearing caps with different colors) allowed the use of a dedicated technique, involving recognizing special color patterns. By doing so, the head of the pedestrians could be identified very precisely. Due to the unfavorable lighting conditions (obstacles casting shadows, etc.), detection was only 100% reliable in a specific detection region.

In the third step, the pedestrians detected in the detection zone in step 2 are tracked, i.e. their locations in the subsequent images are determined. This is achieved by determining the position of the subimage of the pedestrian detected in image *i* in the next image *i*+1 (or image *i*-1, if the pedestrian is tracked backwards). To this end, the so-called *merit function* is minimized. The approach can handle the changes in the shape of the pedestrians due to both the movements of the pedestrians (movement of the arms, head turning, etc.), as well as the position of the pedestrian on the image. The approach determines the position of the center of the position of the pedestrian's head, with an accuracy of one pixel.

Step 4 maps the image coordinates the real-life coordinates. Note that this step also involves remedying the so-called pincushion effect and the rotation of the image. In the fifth step, Kalman filtering is applied to increase the accuracy of the measurement. The final step handles visualization of the results.

In the remainder of the contribution, these steps are described in more detail.

IMAGE CONVERSION AND RADIOMETRIC CORRECTION

Image conversion

First, the still images were gathered from the digital video using a frame grabber, at a frequency of 10 Hz. This results in a long sequence of RGB-images of 720×576 pixels. These images are described by arrays for the Red, Green and Blue values, and contain numbers from 0 to 255 describing the intensity for the respective channels. To reduce the size of the images, the images were cropped (cutting off the parts of the images that are not relevant for a particular experiment).

Radiometric correction

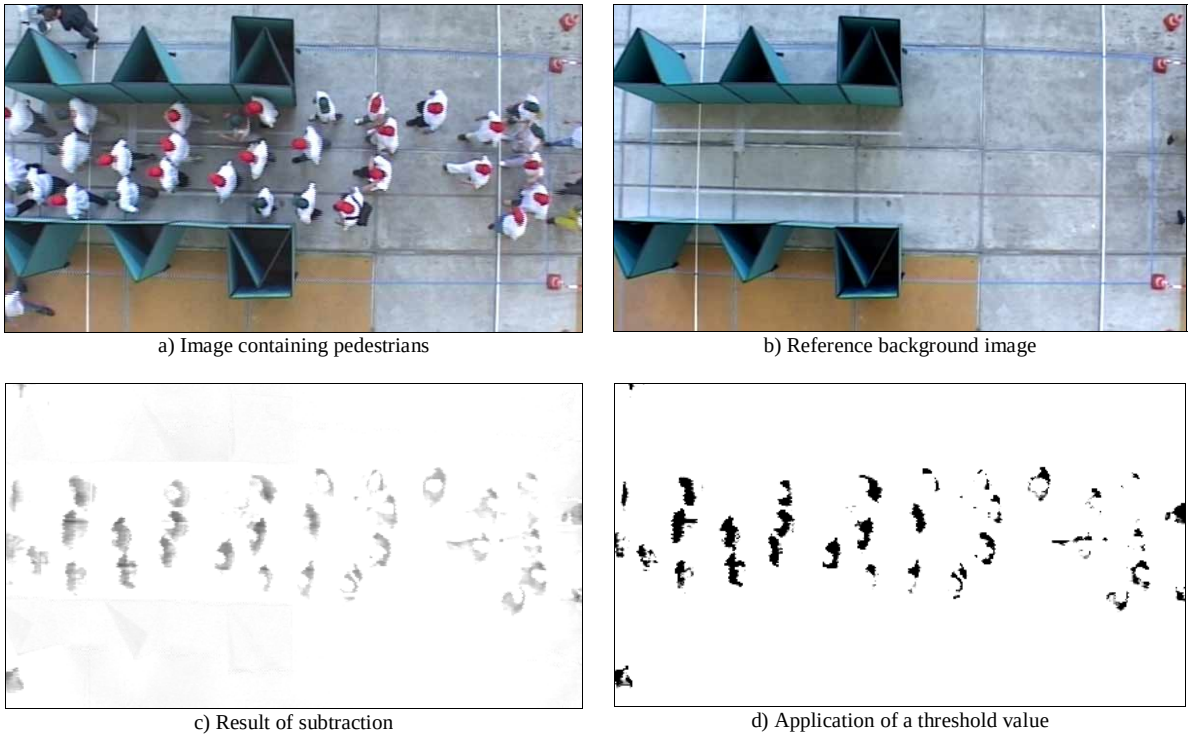
Since due to a variety of reasons ambient conditions were changing during the experiment, some images are brighter than others. Since pedestrian detection is based on identifying differences in the colors of the pixels, it turned out that the pedestrian detection process can be very sensitive to these differences. It was therefore decided to normalize the images using so-called *histogram matching*: by comparing the histograms of the reference image and the current image, the intensities of the current image is adjusted such that they agree with the reference images as much as possible.

PEDESTRIAN DETECTION

The first step of the approach is aimed at identifying the pedestrians from the video images. Several techniques exist (and have been tested) to achieve this step.

One simple approach is to first determine an image of the *empty observed area*, which can used as a reference image. This reference image can then be subtracted from other images that do contain pedestrians, leaving the changes compared to the reference image (i.e. the pedestrians).

The reference image can be determined as follows: for each pixel in an image sequence, the different values of the respective color channels are stored. For these values, the median value is assumed to be the value of the background. Figure 2 b shows the result of this operation. Note that the main assumption is that the probability that the walking surface is empty is larger than the probability that a pedestrian is present, and may therefore be unsuitable in case of dense pedestrian flows.



a) Image containing pedestrians

b) Reference background image

c) Result of subtraction

d) Application of a threshold value

Figure 2 Pedestrian detection by subtracting reference background image (b) from current image (a); results as inverted grayscale image (c), and pixels which intensity values are higher than certain threshold (d).

Figure 2 a-d shows the result of this operation. The pedestrians can be distinguished by identifying the dark spots in the processed images. The pixels marked as pedestrians are subsequently clustered (see next section), representing the pedestrian positions. After some test with different specifications of the difference between the current image and the background image, as well as different threshold values, it was concluded that the approach *is not the most suitable approach for the detection of pedestrians for this particular case*. The main cause is the fact that pedestrians will walk very closely together under dense conditions, and are indistinguishable from one-another. Furthermore, the approach is not directly applicable to distinguish between the two pedestrian groups, based on the color of the caps.

Pedestrian detection and group identification by fuzzy logic

An alternative detection approach was developed for this particular experiment. The approach is based on identification of the colors of the caps (red or a green). To this end, the fuzzy sets ‘red cap’ and ‘green cap’ were constructed by analyzing the color compositions of the caps. The fuzzy sets were represented by trapezoid membership functions, describing the color intensities for the Red, Blue and Green channels for the red caps and the green caps.

Having constructed these fuzzy sets, the total membership to the set ‘red cap’ or the set ‘green cap’ of each pixel on the screen could be determined, describing the *extent in which* a pixel belongs to either set.

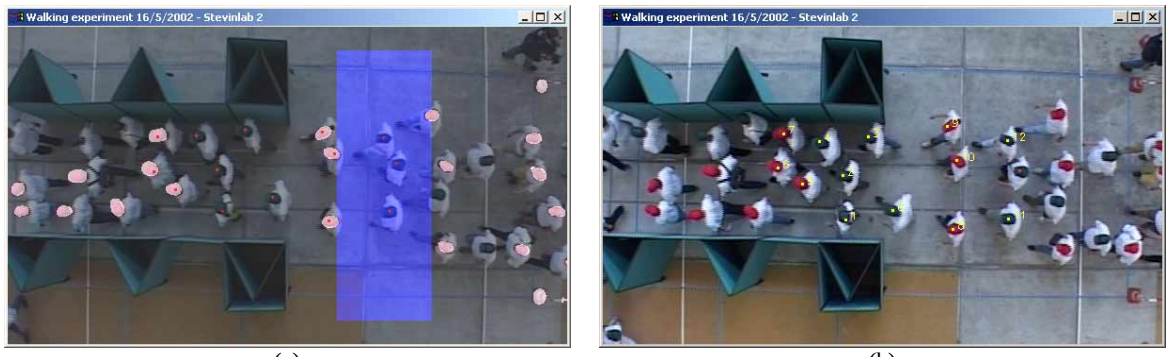
Having performed this step, for each frame (image) and for each pixel on that image, the possibility that the pixel 'belongs' to a pedestrian wearing either a red cap or a green cap has been determined. In the end, each pixel having a membership value higher than $\mu^* = 0.3$ was assumed to belong to a pedestrian. In the remainder, these pixels are referred to as *tagged pixels*.

Clustering pixels to identify pedestrians

When the tagged pixels have been identified, it has to be determined *which* pixel belongs to *which* pedestrian. To this end, different *clustering techniques* have been tested. The most basic technique is based on the Euclidian distance between the tagged pixels, and yielded good results. This is mostly caused by the high quality of the video footage.

When the distance d_{ij} between two tagged pixels r_i and r_j is smaller than some threshold value d_0, i and j are a member of the same cluster C_p (with cluster-mean y_p). When the distance between a cluster C_p and a tagged pixel r_i (defined by the minimum distance between the tagged pixel and any pixel of the cluster) is smaller than the threshold value d_0, pixel p_i belongs to cluster C_p. Finally, when the minimal distance between any pixel of cluster C_p and any pixel of cluster C_q is smaller than the threshold value d_0, C_p and C_q together form a cluster (say, C_p).

The standard clustering technique yielded excellent results in most circumstances, especially when the ambient conditions where good and the color of the caps could be distinguished from the background accurately. However, in several cases, for instance inside the narrow bottleneck (Figure 1a), ambient conditions were unfavorable and the color of the caps (especially the green caps) were not sufficiently distance from other objects in the scene or from the background. As a result, it was decided to detect the pedestrians in a specific area of the image, referred to as the *detection area*, and use a different approach to track the pedestrians from one image to the next (see Figure 3a). It also checks if the pedestrian has already been tracked at an earlier time during the approach.



(a) (b)

Figure 3a) Results from pedestrian detection using fuzzy logic. When an untracked pedestrian is detected, it is tracked using the approach described in the following section. When a pedestrian is successfully tracked, it will no longer be considered during detection. The figure shows detection of 'red cap pedestrians' in the blue detection area. Pedestrians with a red dot already have been tracked; b) Example illustration the pedestrian tracking process.

PEDESTRIAN TRACKING BY OPTIMIZATION

The previous section discussed how pedestrians are detected in the detection zone, and explained the need for an alternative approach for pedestrian tracking. This section presents the developed tracking approach.

Requirements for pedestrian tracking

To track the pedestrians in the subsequent images, the tracking algorithm must be able to:

- Correctly handle changes in the shape of the pedestrian (due to lens distortion, change in perspective, movement of the pedestrian, rotation of the pedestrian, etc.), as well as changes in the color (caused by changing ambient conditions, e.g. when walking through the bottleneck).
- Distinguish between pedestrians that are walking close to each other.

The first requirement implies that the tracking algorithm needs to be adaptive, that is, can respond to (slowly) changing conditions. The second requirement implies that the algorithm needs be able to accuracy distinguish the features of the pedestrians in order to distinguish one pedestrian from the others.

Pedestrian tracking approach

Assume that in image $k = 0$, a pedestrian was detected at location r_k. This location will be the center of a subimage of the pedestrian. This subimage $S_k(r_k)$ has width $2\delta x+1$ and height $2\delta y+1$ (δx and δy are parameters that can be set by the user).

In the next image $k+1$, the area $S_{k+1}(r_k+(sx,sy))$ around r_k is a candidate for the image of the pedestrian in the following image. The likelihood of this being the case is evaluated using the so-called *merit function*:

$$m(sx,sy) = \sum_{i=-\delta x}^{\delta x} \sum_{j=-\delta y}^{\delta y} w(i\delta x, j\delta y) \left| S_{k+1}\left(r_k + \begin{pmatrix} sx \\ sy \end{pmatrix} \right) - S_k(r_k) \right| \quad (1)$$

where sx and sy denote the shift with respect to the location r_k, and where $S_{k+1}(r)$ denotes the subimage around r in image $k+1$; w denotes the weighing function describing that the pixels further from the center of the subimage are of less importance than the pixels near the center. In this case, a Gaussian weighting function has been used

$$w(i\delta x, j\delta y) := \exp\left[-\frac{\sqrt{(i\delta x)^2 + (j\delta y)^2}}{\sigma} \right] \quad (2)$$

for a user-definable factor $\sigma > 0$. To determine the location of the pedestrian in image $k+1$, we determine the shifts sx^* and sy^* that minimize eqn. (1), yielding

$$r_{k+1} = r_k + \begin{pmatrix} sx^* \\ sy^* \end{pmatrix} \quad \text{where} \quad \begin{pmatrix} sx^* \\ sy^* \end{pmatrix} = \arg\min m(sx,sy) \quad (3)$$

Having determined the location of the pedestrian in image $k+1$, we set $k = k + 1$ are reapply the steps. The computations are terminated when the pedestrian walks out of the observed area.

To ensure that the location r_k does not drift from the center of the subimage (i.e. where the cap of the pedestrian is located), for each image k the difference with the first subimage $S_0(r_0)$ is computed as well (assuming that $k = 0$ is the image for which pedestrians have been detected), and added to the merit function with a certain weight $\alpha < 1$, yielding the following definition of the merit function

$$m(sx,sy) = \sum_{i=-\delta x}^{\delta x} \sum_{j=-\delta y}^{\delta y} w(i\delta x, j\delta y) \left| S_{k+1}\left(r_k + \begin{pmatrix} sx \\ sy \end{pmatrix} \right) - (1-\alpha) S_k(r_k) + \alpha S_0(r_0) \right| \quad (4)$$

Finally, since the detection zone is not necessarily located at the edges of the images, the same approach is applied to track the pedestrian back to its starting point (i.e. iterating from image k to k-1).

A tracking-reliability measure e_k is determined by taking the average of the merit function for the duration of the walking trip per unit time. It e_k is high, the merit function had a high value, which is an indication that the tracking results are not reliable. In that case, the tracking is tried again with different parameter settings. If these attempts also fail, the tracking results are discarded.

Figure 3b shows the result of the pedestrian tracking algorithm. The figure depicts pedestrians who have been tracked before (indicated by a yellow dot and a unique id), and the pedestrian that is currently being tracked (indicated by the red dot).

In sum, for each tracked pedestrian, the output of this step consist of

- The unique id of the pedestrian.
- The group to which the pedestrian belongs (red caps or green caps).
- The relevant time instants t_k in seconds.
- The location r_k of the pedestrian at instants t_k in screen pixels.
- A tracking-reliability measure e_k.

The next section briefly discusses how the location measurement are to be converted into real-life coordinates, including how the pincushion effect discussed earlier is to be reduced.

CONVERSION TO TERRESTRIAL COORDINATES

When the pedestrians have been tracked, the image coordinates need to be translated into world coordinates. Furthermore, we need to correct both the rotation of the camera (causing the image to be rotated its center by 1 degree), as well as the pincushion (or lens distortion effect). The rotation of the camera is easily corrected by rotating the tracked locations 1 degree counterclockwise around the center.

Correcting pincushion distortion

Pincushion distortion is a lens effect causing images to be pinched at their center. It is associated with zoom lenses or when adding telephoto adapters and only occurs at the telephoto end of a zoom lens, and is most noticeable when there is a very straight edge near the side of the image frame (see Figure 1).

Several approaches exist to correct for pincushion distortion. Here, the approach of Lin and Fuh (*6*) has been adopted. Their approach is based on estimating the parameters of a distortion model

$$\tilde{r} = \begin{pmatrix} \tilde{r}_1 \\ \tilde{r}_2 \end{pmatrix} = \begin{pmatrix} c_1 & c_2 & c_3 & c_4 \\ c_5 & c_6 & c_7 & c_8 \end{pmatrix} \begin{pmatrix} 1 \\ r_1 \\ r_2 \\ r_1 r_2 \end{pmatrix} \tag{5}$$

using at least eight reference points of which the corrected coordinates are known.

Conversion to real-life coordinates

When the tracked locations are corrected with respect to camera rotation and pincushion distortion, the translation into terrestrial coordinates is straightforward. It is achieved by linear scaling, using the terrestrial coordinates of a number of reference points (i.e. the corners of the walking areas).

FILTERING THE PEDESTRIAN TRAJECTORIES

In the final step of the approach, a Kalman filter is applied to improve the quality of the data as well as to double check the trajectories that have been determined. This section outlines the approach.

Pedestrian kinematics and state dynamics

A Kalman filter combines the predictions of a dynamic model with measurements. For the problem at hand, a very simple dynamic model was used, i.e. the kinematic equations reflecting pedestrian walking behavior. Let r_k denote the position of the pedestrian at instant $t_k = hk$, where h denotes the time stamp. Let v_k denote the velocity of the pedestrian at instant t_k. Assuming that the pedestrian does not change his velocity during the period $[t_k,t_{k+1})$, the location at t_{k+1} equals

$$r_{k+1} = r_k + hv_k \tag{6}$$

For the velocity v_k, we assume the following relation

$$v_{k+1} = v_k + w_k \tag{7}$$

where w_k is a noise-term (describing the unknown acceleration and deceleration of the pedestrian) referred to as *model noise*. In the remainder, we will assume $w_k \sim N(0,Q)$ with $Q = E(w_k w_k')$. Eqns. (6) and (7) can be written in state-space form. Let the *state* of the pedestrian be defined by

$$x_k = (r_k, v_k) \tag{8}$$

Then, the so-called *state-dynamics* are described by the following, linear and time-invariant eqn.

$$x_{k+1} = \begin{pmatrix} I & hI \\ 0 & I \end{pmatrix} x_k + w_k = Ax_k + w_k \tag{9}$$

where I denotes the unit matrix.

Measurement equations

The so-called *measurement equation* relates the state x_k at instant t_k to measurement y_k collected at the same instant t_k. In this case, the following linear time-invariant measurement equation was used

$$y_k = C_k x_k + \varepsilon_k = \begin{pmatrix} I & 0 \end{pmatrix} \begin{pmatrix} r_k \\ v_k \end{pmatrix} + \varepsilon_k \tag{10}$$

where $\varepsilon_k \sim N(0, R)$ with $R = E(\varepsilon_k \varepsilon_k')$ reflect the so-called measurement noise, describing the measurement errors made during data collection. Eqn. (10) shows that the positions r_k are measured with a certain error (i.e. $y_k = r_k + \varepsilon_k$).

Kalman filter equations

A Kalman filter generally consist of two steps: prediction and correction. During prediction, the state $x_{k|k}$ of the system is updated using the state dynamic equation (9). Besides predicting the state, the covariance $P_{k+1|k}$ of the prediction is computed. This leads to the following prediction equations

$$\begin{aligned} x_{k+1|k} &= A x_{k|k} \\ P_{k+1|k} &= A P_{k|k} A' + Q \end{aligned} \tag{11}$$

subject to

$$x_{0|0} = x_0 \quad \text{and} \quad P_{0|0} = P_0 \tag{12}$$

where x_0 is the initial state estimate and P_0 is the covariance of this estimate.

The second step of the Kalman filter pertains to correcting the predictions using the available measurements. In this correction step, the measurements are used as follows

$$\begin{aligned} x_{k+1|k+1} &= (I - K_k C) x_{k+1|k} + K_k y_{k+1} \\ P_{k+1|k+1} &= (I - K_k C) P_{k+1|k} (I - K_k C)' + K_k R K_k' \end{aligned} \tag{13}$$

where the so-called *filter-gain* K_k equals

$$K_k = P_{k+1|k} C' (C P_{k+1|k} + R)^{-1} \tag{14}$$

To determine the reliability of the measurements, the absolute difference between the predicted location and the measurement was computed and stored. The filter will warn the user of the system when this difference becomes too large, which is likely to be caused by incorrect tracking.

APPLICATION RESULTS

After fine-tuning the algorithm, all pedestrians participating in the experiment could be detected and tracked. Occasionally, false detections were encountered, e.g. when a pedestrian with green colored pants entered the detection zone. During tracking, these false detections were recognized by high values of the merit function, caused by the fact that the falsely detected object could not be found in the subsequent images. As a result, false detections (amounted to approximately 1.5% of the total number of pedestrians) were easily discarded from the data. Visual verification also revealed that in approximately 1% of the cases, tracking of the pedestrian could not be completed successfully. In these cases, the trajectories have been determined by alternative methods (e.g. by combining detection and tracking, or manually).

Results for two bottleneck experiments

The main objective of this paper is to present the approach to derive microscopic pedestrian data from (digital) video footage. Nevertheless, this section discusses some new insights into microscopic pedestrians behavior. The results discussed here are based on the bottleneck experiments, i.e. the *narrow bottleneck* (width of 1 m) and the *wide bottleneck* (width of 2 m).

Figure 4 shows different projections of seven trajectories (wide bottleneck). Using these projections, different microscopic processes can be identified. For instance, the crossing of trajectories in the *tx*-plane indicates a fast pedestrian overtaking a slower pedestrian. Let us now briefly discuss some interesting results that can be observed from the trajectories. For a more in-depth discussion, we refer to (*5*).

Oscillations and step frequency

Note the oscillations in the trajectories that can be observed in the *ty*-plane and in the *xy*-plane. These oscillations show that while walking, pedestrians move from side to side, especially in the region upstream of the bottleneck. Inside the bottleneck, these oscillations are less apparent. Further investigation has shown

that the frequency of these oscillations is approximately 2 Hz, which reflects the step frequency of the pedestrians. This number is in line with the results reported by Weidmann (*7*).

The amplitude of these oscillations is somewhere between 10 to 20 cm, and depends on the walking speed of the pedestrians: when the speed is low (congestion), the amplitude increases, when the speed is high (free flow), it decreases. The frequency of the oscillations is less sensitive to the speed: only for very low speeds, the frequency reduces under 2 Hz.

Lane formation and following behavior inside bottlenecks

The trajectories also reveal how the capacity of the bottleneck depends on the behavior of the pedestrians: the narrow bottleneck case clearly shows how the pedestrian stream inside the bottleneck is effectively divided into two lanes (see Figure 5). These lanes appear to be 'zipped': pedestrian effectively follow the pedestrian straight ahead at a median time-headway of approximately 1.22 s, while paying little attention to the pedestrian diagonally in front (although the latter is nearer in terms of distance). The median time headway of the collected at the total cross-section is approximately 0.61 s, yielding a capacity of 1/(1.0*0.61) = 1.6 P/s/m for a bottleneck of 1 m.

This result suggests that the when the bottleneck becomes too narrow – say 80 cm – and only a single lane can be formed, the capacity reduces substantially to 1/(0.8*1.22) = 1.02 P/s/m. We thus argue that the capacity depends discontinuously on the bottleneck width.

Congestion and adaptive behavior

A final interesting observation is the adaptive behavior of pedestrians (narrow bottleneck case) during congestion. It is clear that the amount of space needed by a pedestrian depends on the speed of the pedestrians (*7*). However, from the experiments it was concluded that this space also depends on the walking distance to the actual bottleneck. When pedestrians are far away from the bottleneck, they tend to take up a relative large amount of space given their speed: apparently, the pedestrians do consider to acquire personal benefits from standing very close to the other pedestrians in their direct neighborhood. When approaching the bottleneck, the speed remains approximately the same but the density (and thus the required space) increases: when pedestrians approach the bottleneck, they will start 'pushing' in order to increase their personal benefit (travel time, probability of being quick in passing the bottleneck).

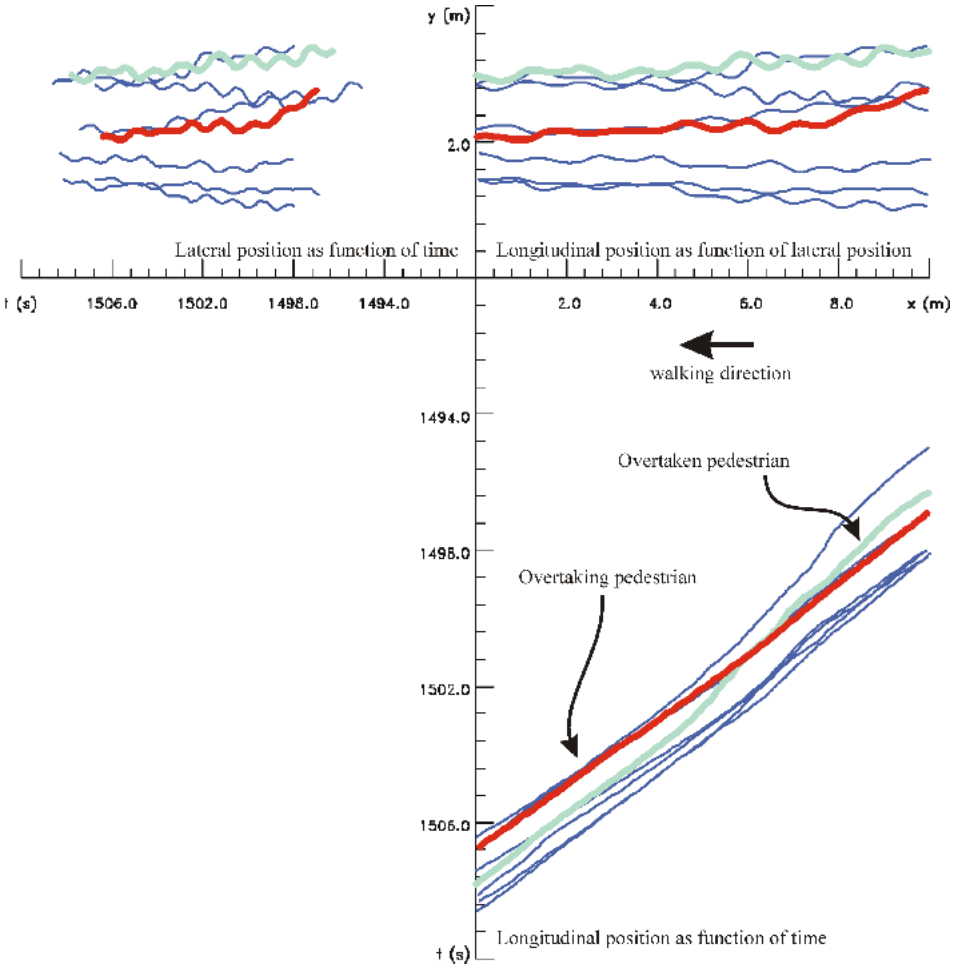


Figure 4 Results from application of the algorithm to footage of the wide bottleneck experiment. The figure shows different projects of the trajectories on the respective planes (*xy*-plane, *tx*-plane, and the *ty*-plane).

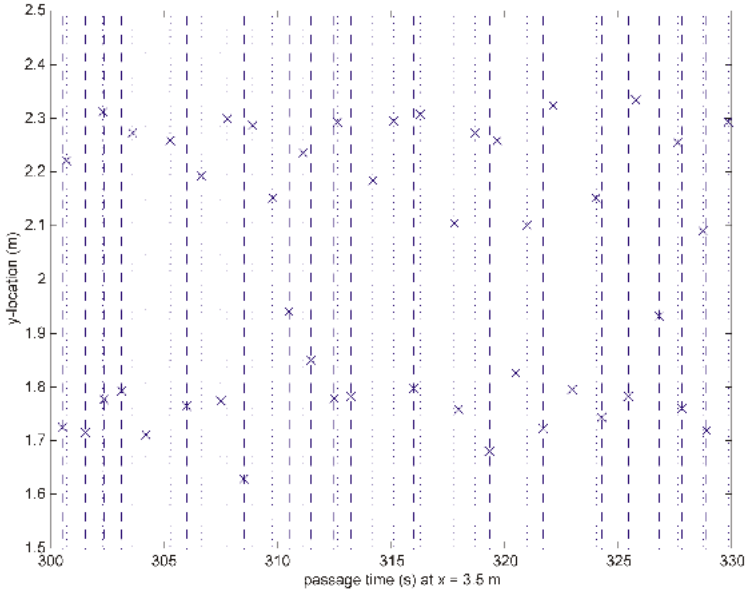


Figure 5 Lateral positions of pedestrians passing the cross-section at x = 3.5 m for congested period [300s, 330s]. Figure clearly shows how sample can be divided into pedestrians with y < 2m and y > 2m while passing.

CONCLUSIONS AND OUTLOOK

This contribution describes a new approach to determine microscopic pedestrian data from digital video measurements. The approach consists of six steps, including pedestrian detection and tracking, and was applied to data collected during walking experiments conducted at Delft University of Technology. It was found that the approach is suitable to detect all pedestrians in the detection zone, and track them with a high accuracy. False detections are identified by two performance measures.

The newly developed detection and tracking techniques can also be applied to other types of video footage (using alternative detection techniques) and under different circumstances. For instance, the technique may be applied to real-life situations, e.g. by measuring from a high building, or another high vantage point. Depending on the position of the camera, additional photogrammetric techniques need to be applied before detection and tracking of the pedestrians becomes possible (e.g. orthorectification). We belief that the tracking technique developed here can be applied to other data sources without much modification. The automated detection of pedestrians may however be more involved in real-life situations, despite the fact that first results on real-life data are promising.

The results of the approach are used to improve our knowledge of the fundamental properties of walking behavior. As such, the collected data are either used to develop new theories and models, or will be used to calibrate and validate existing walker models, such as the NOMAD model (*2*).

Acknowledgements - This research is partly funded by the Social Science Research Council (MaGW) of the Netherlands Organization for Scientific Research (NWO). The authors wish to thank the anonymous reviewers for their constructive criticisms and suggestions.

REFERENCES

1. Helbing, D., I. Farkas, and T. Vicsek (2000) Simulating Dynamical Features of Escape Panic. *Nature 407*, 487-490.
2. Hoogendoorn, S.P. and P.H.L. Bovy (2002). Normative Pedestrian Behaviour Theory and Modelling. *15th International Symposium on Transportation and Traffic Theory*, Adelaide, Australia 2002.
3. Blue, J.B. and J.L. Adler (2001) Cellular automata microsimulation for modeling bi-directional pedestrian walkways. *Transportation Research B 35*, pp. 293-312.
4. Brackstone, M. and M. McDonald (1995). *The microscopic modelling of traffic flow: weaknesses and potential developments*. Traffic and Granular Flow.
5. Daamen, W. and S.P. Hoogendoorn (2002). *Experimental Research of Pedestrian Walking Behavior*. Submitted to the 2003 Annual Meeting at the Transportation Research Board.
6. Lin, Y.-C., and C.-S. Fuh (2002). *Correcting Distortion for Digital Cameras*. Proceedings of the National Scientific Council ROC(A). Vol. 24, No. 2., 115-119.
7. Weidmann, U. (1993) *Transporttechnik der Fussgaenger*. ETH Zuerich, Schriftenreihe IVT-Berichte 90, Zuerich (In German).
8. Goffman, E. (1971). *Relations in Public: Microstudies in the Public Order*. New York. Basic Books.
9. Sobel, R.S., and N. Lillith (1975). Determinant of Nonstationary Personal Space Invasion. *Journal of Social Psychology*. Vol. 97, 39-45.
10. Dabbs, J.M., and N.A. Stokes (1975). Beauty is Power: the Use of Space on the Sidewalk. *Sociometry*. Vol. 38, no. 4, 551-557.
11. Willis *et al.* (1979). Stepping aside: correlates of Displacements in Pedestrians. *Journal of Communication*. Vol. 29, no 4., 34-39
12. Morrall, J.F., L.L. Ranayaka, and P.N. Seniviratne (1991). Comparison of Central Bussines District Pedestrian Characteristics in Canada and Sri-Lanka, *Transportation Research Record* 1294, 57-61.
13. Lam, W.H.K., Y.S. Jodie, Y.S. Lee, and K.S. Chan (2002). A Study of the Bi-Directional Flow Characteristics at Hong-Kong Signalized Walking Facilities. *Transportation 29*, 169-192.
14. Lam, W.H.K., and C.Y. Cheung (2000). Pedestrian Speed-Flow Relationships for walking facilities in Hong-Kong. *Journal of Transportation Engineering, ASCE*, 126(4), 343-349.
15. Hughes, R. Pedestrian flows in Mekka.
16. Willis, A., R. Kukla, J. Kerridge, and J. Hine (2001). Laying the Foundations: the Use of Video Footage to Explore Pedestrian Dynamics in PEDFLOW. In: *Pedestrian and Evacuation Dynamics*, Springer, 181-186.

17. Wang, D. (1998) Unsupervised Video segmentation Based On Watersheds And Temporal Tracking, *IEEE Trans. Circuits Syst. Video Technol.*, 8(5), 539-546
18. Foresti, G. L. (1999) Object Recognition And Tracking For Remote Video Surveillance, *IEEE Trans. Circuits Syst. Video Technol.*, 9(7), 1045-1062
19. Salembier, P. , F. Marqu s, M. Pard s, J. R. Morros, I. Corset, S. Jeannin, L. Bouchard, F. Meyer, B. Marcotegui (1997) Segmentation-Based Video Coding System Allowing The Manipulation Of Objects, *IEEE Trans. Circuits Syst. Video Technol.*, 7(1), 60-74
20. A. J. Lipton, H. Fujiyoshi, R. S. Patil (1998) Moving Target Classification And Tracking From Real-time Video, *Applications of Computer Vision, 1998. WACV '98. Proceedings., Fourth IEEE Workshop on*, pp. 8-14
21. Wang, Y., R.E. Van Dyck, J. F. Doherty (2000) Tracking Moving Objects in Video Sequences, *Proc. Conference on Information Sciences and Systems*, Princeton, NJ