Automatic Detection of Dangerous Motion Behavior in Human Crowds

Barbara Krausz, Christian Bauckhage
Fraunhofer IAIS
53754 Sankt Augustin, Germany
{barbara.krausz, christian.bauckhage}@iais.fraunhofer.de

Abstract

Tragically, mass gatherings such as music festivals, sports events or pilgrimage quite often end in terrible crowd disasters with many victims. In the past, research focused on developing physical models that model human behavior in order to simulate pedestrian flows and to identify potentially hazardous locations. However, no automatic systems for detection of dangerous motion behavior in crowds exist.

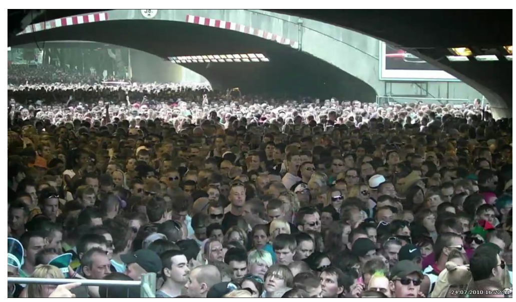
In this paper, we present an automatic system for the detection and early warning of dangerous situations during mass events. It is based on optical flow computations and detects patterns of crowd motion that are characteristic for hazardous congestions. By applying an online change-point detection algorithm, the system is capable of identifying changes in pedestrian flow and thus alarms security personnel to take necessary actions.

1. Introduction

Mass events are (and always have been) popular in human societies all over the world. Nowadays, typical examples include sports events, festivals, or concerts. Due to increasing populations and higher mobility, mass events attract ever growing numbers of visitors and adequate safety measures are becoming more and more important. Nevertheless, despite of all precautions and the use of technology such as video surveillance, deadly stampedes and crowd disasters still occur rather frequently (see Table 1).

At such mass gatherings, the density of the crowd easily becomes extremely high; studies report densities up to 10 people per square meter [20]. High pedestrian densities usually come along with typical patterns of mass behavior such as *stop-and-go waves* or *crowd turbulences* [8]. Essentially, stop-and-go waves show alternating forward pedestrian motion and backward gap propagation. They occur when the pedestrian density is critically high and unobstructed pedestrian flow becomes impossible [9]. In even more densely packed crowds, people moving involuntarily induce sudden movements of other people nearby. These crowd turbulences propagate through the crowd causing people to stumble and to fall down. As a result, most people die by suffocating due to dangerous pressure of up to 4500 N/m on their chests [8].

(a)



(b)

Figure 1. Views of the festival area of the Loveparade 2010 in Duisburg, Germany [15].

Work towards the prevention of crowd disasters includes simulations of human behavior in crowds. Using such simulations, one can identify and defuse places in an environment that are potentially dangerous. However, human reactions and mass behavior are often unpredictable, in particular, whenever alcohol consumption is an integral part of a mass event. Moreover, gathering areas are often large and become unmanageable if there are many thousands of visitors. In cases like these, automatic video surveillance systems may help to estimate the visitor density and to detect indicators of critical situations. This can buy crucial time in

Year	Place	Deaths
2010	Loveparade, Duisburg	21
2010	Water Festival, Phnom Phen	>380
2006	Stadium, Yemen	51
2006	Pilgrimage, Mena	363
2005	Religious Procession, Bagdad	>640
1999	Subway Station, Minsk	53
1990	Pilgrimage, Mena	1426
1989	Stadium, Sheffield	96

Table 1. Examples of recent deadly stampedes and crowd disasters (see [8]).

which security personnel can be dispatched and streams of pedestrians can be redirected.

In this paper, we develop a system for motion behavior analysis of masses that computes dense optical flow fields which can be calculated in real-time. By using optical flow, our system avoids the need for detection and tracking of individual pedestrians which is often impossible due to inappropriate camera viewpoints or occlusions due to the large number of people. Based on histograms of optical flow, we propose methods to automatically detect congestions and shock waves. We test this approach on video footage from the crowd disaster at the Loveparade 2010 in Duisburg, Germany. To the best of our knowledge, this is the first vision-based system that robustly detects dangerous situations like overcrowding and crowd turbulences in real-time.

2. Related Work

Pedestrian dynamics have been studied intensively for more than 40 years. Recently, knowledge about crowd dynamics has been used to improve evacuation strategies in emergency situations and to prevent congestions and overcrowding (see [22] for an overview). Simulations are a standard tool in the study of self-organizing effects of large groups of pedestrians. Physical models modeling pedestrians based on the analogy to gases, fluids or granulates have been developed in order to account for individual behavior. The *social force model* [10, 7] as well as *cellular automata* [4, 13] which both model pedestrian dynamics on a microscopic level are among the more widely used approaches.

In addition to simulations, experimental studies are conducted in order to understand human behavior and improve existing physical models. Parameters such as crowd density, speed, flow, and crowd pressure (see [9, 19] for definitions) are determined either manually [18] or by means of digital image processing [14, 12]. Usually the resulting representations are based on experimental data and do not consider real data. Moreover, video-based experiments are typically carried out using top-view cameras in order to avoid occlusions and to facilitate automatic video analysis. Techniques that are applied in this context usually detect and track individuals but there also exist holistic approaches that make use of optical flow features.

Over the years, various visual tracking approaches have been reported that were specifically developed for tracking pedestrians in crowded scenes [23, 3, 21]. More recently, ideas adopted from simulations of pedestrian dynamics were incorporated into the design of visual tracking systems. Ali and Shah [2] present a tracking framework inspired by the cellular automaton model [4]. They automatically calculate force fields that integrate information on human behavior as well as the locations of obstacles and important regions such as exit doors. In their previous work [1], Ali and Shah propose a flow segmentation framework which enables them to detect changes in flow patterns. Mehran *et al*. [16] adopt ideas from the social force model and estimate interaction forces in order to detect abnormal events. All these works do not detect and track individuals. Instead, they apply the technique of *particle advection* that places particles onto a grid and moves them according to the underlying optical flow field. However, in our case, particle advection is not applicable due to inappropriate camera view points and the resulting occlusions in situations of high pedestrian density.

3. Detection of Dangerous Mass Behavior

Traditional approaches to real-time surveillance detect and track individual people. However, in surveillance of mass gatherings this is not feasible, since hundreds of people are visible from a camera's viewpoint which necessitates considerable computational efforts. Instead, we consider dense optical flow fields to determine major motion patterns and motion directions in the crowd. In addition to computational efficiency, this also guarantees the privacy of people being monitored.

We compute dense optical flow fields using the method proposed by Farnebäck [6]. Herein, quadratic polynomials are used to estimate translations of a local neighborhood and motion vectors are determined from polynomial expansion coefficients. In order to account for spatially varying motion patterns, we superimpose a grid of cells over the video frames. In case of camera viewpoints similar to that in Figure 1(b), grid cells towards the back of the scene are smaller taking perspective distortions into consideration. Given the dense optical flow field for a grid cell, we compute two-dimensional histograms ($36 \cdot 100$ bins) of motion magnitude and motion direction of the flow vectors.

3.1. Congestion Detection

In congested areas, the crowd is moving slower or has come to a halt. When observing such a situation from above, one can estimate the people's velocities by considering the magnitude of optical flow vectors. However, when the camera's viewpoint is similar to the viewpoint in Figure 1(b) where people are going towards the camera, it is not straightforward to estimate their velocities due to perspective effects. In this case, we make use of the following observation: In a congested area, one can observe that people go very slowly while stepping from one foot to the other in order to keep their balance resulting in oscillating motions. In fact, Liu *et al*. [14] reported experiments with several groups of pedestrians moving with different speeds from 0.26 m/s to 1.72 m/s. Their movements were filmed from above and the authors generally observed lateral oscillation in the trajectories. This is due to the fact that people do not move along a straight line, instead, it is a characteristic of human gait, that they tend to swing laterally. Liu et al. [14] also observed that while the amplitude of the lateral oscillation is higher for lower speeds, the frequency increases for higher speeds. The authors found linear relationships between the velocity and the amplitude as well as between the velocity and the frequency.

Given these observations, we propose a method for the automatic detection of dangerous congestion situations. First, we compute dense optical flow and corresponding two-dimensional histograms of motion direction and magnitude. Then, we average the histogram over a short time interval. When observing a scene from above, we can directly make use of the magnitude of optical flow motion vectors by computing the center of mass $(c_{dir,t}, c_{mag,t})$ of the histogram and taking $c_{mag,t}$ as a feature for congestion detection for topview cameras. Low values of $c_{mag,t}$ indicate that the velocity of the people is low which might be indicative for a congestion. Contrarily, when the scene is observed by a front view camera (see Figure 1(b)), we make use of increasing lateral oscillations in congestions. Here, histograms that are indicative of congestion situations show motion along two major directions (rightwards and leftwards) which reflect lateral oscillation of the people's upper bodies. Such histograms show a high degree of symmetry (see Figure 2) so that we measure the mirror symmetry of an optical flow histogram and consider the resulting value a feature for congestion detection for front view cameras.

We compute the symmetry measure by summing the absolute differences between the histogram and a flipped version of itself. As described above, we subdivide each frame into a set of cells with cells in the background of the scene being smaller to account for effects of viewing perspective. Let $H_{i,c}(dir, magn)$ be the two-dimensional histogram of direction and magnitude of cell c at time i. Then, denoting by $\hat{H}_{i,c}(dir, magn)$ the flipped version of $H_{i,c}(dir, magn)$, we compute

$$sym_{i,c} = \sum_{dir,mag} \left| \hat{H}_{i,c}(dir, magn) - H_{i,c}(dir, magn) \right|. \quad (1)$$

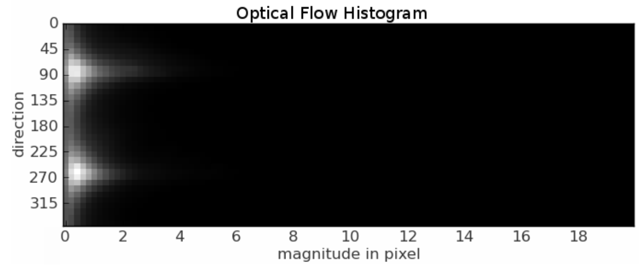


Figure 2. A histogram of optical flow that is characteristic for motion in congestion situations. It shows small motion along two major directions. This left- and rightward motion is caused by people swinging laterally to keep their balance.

Accordingly, low values of $sym_{i,c}$ indicate that $H_{i,c}(dir, magn)$ is highly mirror-symmetric and is indicative for a congested area. Figure 3 shows results obtained from video footage of the Loveparade.

Now, we apply a sequential change-point detection algorithm for detecting unusual events and congestions in particular. The method proposed by De Oca *et al*. [5] extends the conventional cusum algorithm [17]. It is a non-parametric cusum algorithm that allows for distributions varying in time and uses historical data for obtaining suitable thresholds above which an alarm is raised. We extend this algorithm to compute an additional measure that characterizes the severity of an alarm.

Let us denote an observation $sym_{t,c}$ as Y_i and consider a sequence of observations $\{Y_i\}_{i=1}^N$. We use previous observations $\{Y_j\}_{j=i-k-l}^{i-k}$ to estimate a reference distribution where k is a fixed time interval and l is a fixed number of historical observations that are used for estimating the reference distribution. Next, we denote the upper and lower α-percentiles of the reference distribution as $Q(\alpha)$ and $Q(1-\alpha)$, respectively, where α is specified by the user and controls the degree a deviation from the reference distribution is considered as critical. The cusum algorithm continuously accumulates deviations of incoming observations from the reference distribution:

$$\begin{aligned} S_i^+ &= \max\{0, S_{i-1}^+ + Y_i - Q(\alpha)\}, S_0^+ = 0 \\ S_i^- &= \max\{0, S_{i-1}^- + Q(1-\alpha) - Y_i\}, S_0^- = 0 \end{aligned} \quad (2)$$

It raises an alarm if either $S_i^+ > \Theta$ or $S_i^- > \Theta$, where in the first case, we detect an upward shift of the signal and a downward shift in the latter case. The threshold Θ is calculated from the reference distribution as follows: Suppose that the sequence of observations is drawn from the reference distribution, that is, no anomaly occurs. Using a bootstrap resampling method, Θ is selected so that the probability of a false alarm is equal to γ, a parameter specified by the user. For that purpose, M sequences are sampled from the reference distribution. For each sampled sequence

m, cusum statistics according to equation 2 are computed and $\max\{S^+_{sampled,m}, S^-_{sampled,m}\}$ is determined. Next, for each sampled sequence m, we select the maximum value of $\max\{S^+_{sampled,m}, S^-_{sampled,m}\}$ and compute the threshold Θ as the $(1-\gamma)$-percentile from these maximum values.

Whenever either S^+_i or S^-_i exceeds the computed threshold Θ, we raise an alarm. De Oca *et al.* [5] also propose a method for detecting the end time of an alarm. They apply a slope testing technique for detecting a downward trend in the cusum statistics which indicates that the deviations from the reference distribution become smaller: Without loss of generality, we assume that S^+_i exceeds Θ at time a (The same rationale holds for S^-_i.). Then, a linear regression model is continuously fitted to a sliding window of cusum values $\{S^+_i\}^n_{i=n-\nu+1}$ for $n = a, a+1, \ldots$ and ν being a fixed size of the sliding window. The end time of an alarm is detected, when the slope of the linear regression model is less than or equal to zero. Then, cusum statistics S^+_i or S^-_i, respectively, are set to zero. Additionally, we propose to measure the severity of the raised alarm as a value $L \in [0 \ldots 1]$ by computing the angle of the regression line in degrees and dividing it by 90°. This is motivated by idea that the slope of the linear regression model of the cusum statistics S^+_i (or S^-_i, respectively) depends on the deviation of the current observation Y_i to the reference distribution: The higher the deviation is compared to the reference distribution, the larger the slope of the regression line is. If L is near to one, the slope is large and the situation is considered to be very critical.

In particular, congestions are characterized by low values of $sym_{t,c}$ as described above. Thus, an alarm raised by the system is very severe and may indicate a congestion, if L is near to one and S^-_i exceeds the threshold indicating that $sym_{t,c}$ decreases due to optical flow histograms becoming more and more mirror-symmetric. Section 4 gives results obtained from video sequences showing congested areas.

3.2. Detection of Crowd Turbulences

In areas of extremely high pedestrian density, the movement of a person affects other nearby people. Shock waves might occur and propagate through the crowd. Situation like these are extremely dangerous since people cannot control their motion anymore but are moved by the crowd; people who loose their balance and fall down in a shock wave typically get crushed and suffocate.

Shock waves are characterized by a sudden increase of the magnitude of the optical flow motion vectors. Moreover, since several people in a local neighborhood move into the same direction, the standard deviation of local motion directions σ_{dir} is small. Therefore, in order to accomplish the automatic detection of shock waves, we divide the frame into C cells. For each cell c and each time t, we compute $\mu_{t,c}$, the average magnitude of optical flow motion vectors. Then, we compare the current average magnitude to previously observed values. Let

$$\mu_{prev,c} = \frac{1}{r} \sum_{i=t-k}^{t-1} \mu_{i,c},\ r > 1 \tag{3}$$

be the average magnitude value of frames $t-r$ to $t-1$ in cell c and $\mu_{t,c}$ the average magnitude of the current frame t in cell c. Then, we compute a measure for the increment of local velocities as $a_c = \frac{\mu_{t,c}}{\mu_{prev,c}}$. As shock waves are characterized by a sudden acceleration of the crowd and a small standard deviation of motion directions, we also compute a value $p_c = \frac{a_c}{\sigma_{dir,c}}$ for each cell c. In those cells of a frame where a shock wave is observed to propagate, p_c will be high. Thus, we compute the average value of all p_c for the entire frame

$$\mu_{frame} = \frac{1}{C} \sum_{c=0}^{C} p_c,\ C = \text{total number of cells} \tag{4}$$

as well as the standard deviation σ_{frame} of all p_c values. This allows us to detect shock waves in locations where $|p_c - \mu_{frame}|$ is greater than $n \cdot \sigma_{frame}$, $n = 2.5$. Figure 4 shows an example of automatically detected shock wave regions colored in red.

4. Results

We tested our approach for detecting critical situations in crowds on video footage from the crowd disaster at the Loveparade 2010 in Duisburg, Germany. In this terrible stampede, 21 visitors died and more than 500 were injured. The festival area was monitored by seven cameras where three of them were static cameras. Video footage recorded by all cameras can be downloaded from [15]. In particular, we analyzed motion patterns from camera 15 that was located in the western bridge area (see Figure 1(b) for an exemplary screen-shot). Figure 3 shows the development of $sym_{t,c}$ measuring the mirror symmetry of the optical flow histograms. To create this plot, we averaged histograms of optical flow over a time period of 10 seconds and summed the values of $sym_{t,c}$ for the cells in the scene foreground. We automatically detect change-points in an online manner using the cusum algorithm presented in section 3.1. Here, we use $l = 90$ observations of historical data to estimate a reference distribution which corresponds to a time interval of 15 minutes whereas k is set to 30 observations (= 5 minutes). We set α, the parameter to control the degree a deviation from the reference distribution is considered as critical, to 0.95 and γ which controls the probability of false alarms to 0.1. Next, the parameter ν used in the detection of the end time of the alarm is set to 8 and M specifying the number of sampled sequences for computing a suitable threshold is set to 100. Alarms that have been raised by our system are colored in red, if the signal is at a low level, whereas a jump to a high value is marked in green. In the lower part, the alarm level L is depicted which measures the severity of the alarm.

Comparing the automatically detected alarms with the video footage reveals that seven out of ten alarms correspond to anomalies in the video, e.g. ambulances or police cars crossing the scene. Figure 3 gives interpretations for these alarms. In particular, at about 16:27 the system raises a severe alarm ($L > 0.7$) and reports low values of $sym_{t,c}$ which is indicative for a congestion. In fact, at this time point the crowd is densely packed and has come to a halt. In this situation, our system would have detected a very critical situation and alarmed the security personnel to take necessary actions in order to prevent a deadly stampede.

Only two false alarms at the beginning of the video recordings are reported which can be explained by the fact that our system has not yet integrated enough observations of normal crowd behavior. A third false positive alarm is raised at 16:17, but has a low severity ($L = 0.1$) and lasts for just a few seconds.

We also tested our approach on a dataset recorded by the Hermes project [11] under laboratory conditions. Here, pedestrians walk through a corridor with a bottleneck at the end of the corridor. The scene is recorded by two topview cameras which distinguished this dataset significantly from the Loveparade videos. We compute the center of mass $(c_{dir,t}, c_{mag,t})$ of the histogram and take $c_{mag,t}$ as a feature for congestion detection for topview cameras. Using the proposed change-point detection method, our system successfully detects different phases: Firstly, it detects that people enter the field of view. Next, it recognizes a phase where the corridor is congested and then it correctly identifies the timepoint when people leave the camera's field of view. However, the system reports a false alarm at the beginning, because it has not yet integrated enough observations.

We also tested our approach for detecting shock waves on videos from the Loveparade stampede. Particulary, camera 13 (see Figure 1(a)) monitoring the main entry ramp to the festival area shows short sequences of shock waves propagating through the densely packed crowd. Regions of high pressure are automatically detected by the method presented in section 3.2 and are depicted in Figure 4.

Figure 4. Detection of Crowd Turbulences. Camera 13 which monitored the main entry ramp to the Loveparade festival area shows short sequences of shock waves that are propagated. We automatically detect regions of high pressure which are colored red in the above Figure.

5. Conclusions

Mass events are becoming more and more popular. Despite of more than four decades of research in crowd dynamics and pedestrian simulation for improving security, terrible stampedes such as the crowd disaster at the Loveparade 2010 occur all over the world rather frequently.

We presented a system that automatically detects critical situations in crowded scenes and warns security personnel in order to take all necessary actions to prevent a crowd disaster. Apparently, our computationally efficient approach to motion pattern analysis is applicable in many different situations, for example at concerts, in stadiums or in subway stations. The described features of motion vector histograms reflects a very general pattern of crowd motion that is characteristic for hazardous congestions. Moreover, it preserves privacy of people being monitored at the mass event, since it does not detect and track individual people.

Acknowledgements This work has been supported by the EU funded Research Project PRONTO (FP7-ICT 231738).

References

[1] S. Ali and M. Shah. A lagrangian particle dynamics approach for crowd flow segmentation and stability analysis. In *CVPR*, 2007.

[2] S. Ali and M. Shah. Floor fields for tracking in high density crowd scenes. In *ECCV*, 2008.

[3] G. Brostow and R. Cipolla. Unsupervised bayesian detection of independent motion in crowds. In *CVPR*, 2006.

[4] C. Burstedde, K. Klauck, A. Schadschneider, and J. Zittartz. Simulation of pedestrian dynamics using a 2-dimensional cellular automaton. *Physica A*, 295(3–4):507–525, 2001.

[5] V. M. De Oca, D. Jeske, Q. Zhang, C. Rendon, and M. Marvasti. A cusum change-point detection algorithm for non-stationary sequences with application to data network surveillance. *J. of Systems and Software*, 83:1288–1297, 2010.

[6] G. Farnebäck. Two-frame motion estimation based on polynomial expansion. In *SCIA*, 2003.

[7] D. Helbing, I. Farkas, and T. Vicsek. Simulating dynamical features of escape panic. *Nature*, 407(6803):487–490, 2000.

Figure 3. Detection of anomalies in the pedestrian flow at the Loveparade 2010 recorded by camera 15. Alarms raised by our system are colored in red for a abnormally decreasing signal, whereas jumps to a high level are depicted in green. In the lower part, the computed severity $L \in [0 \ldots 1]$ of the alarm is shown. Clearly, our system has detected abnormal events that deviate from typical observations. In particular, our system raises a severe alarm at about 16:27 when a the crowd is jamming up.

[8] D. Helbing and A. Johansson. Pedestrian, crowd and evacuation dynamics. In R. A. Meyers, editor, *Encyclopedia of Complexity and Systems Science*. Springer, 2009.

[9] D. Helbing, A. Johansson, and H. Z. Al-Abideen. Dynamics of crowd disasters: An empirical study. *Physical Review E*, 75(4):046109–1–7, 2007.

[10] D. Helbing and M. P. Social force model for pedestrian dynamics. *Physical Review E*, 51(5):4282–4286, 1995.

[11] Hermes. BMBF Research Programme. http://www2.fz-juelich.de/jsc/appliedmath/ped/projects/hermes, 2011.

[12] A. Johansson, D. Helbing, H. Z. Al-Abideen, and S. Al-Bosta. From crowd dynamics to crowd safety: A video-based analysis. *Advances in Complex Systems*, 11(04):497–527, 2008.

[13] A. Kirchner and A. Schadschneider. Simulation of evacuation processes using a bionics-inspired cellular automaton model for pedestrian dynamics. *Physica A*, 312(1-2):260–276, 2002.

[14] X. Liu, W. Song, and J. Zhang. Extraction and quantitative analysis of microscopic evacuation characteristics based on digital image processing. *Physica A*, 388(13):2717–2726, 2009.

[15] Lopavent. Dokumentation Loveparade. www.dokumentation-loveparade.com, 2010.

[16] R. Mehran, A. Oyama, and M. Shah. Abnormal crowd behavior detection using social force model. In *CVPR*, 2009.

[17] E. Page. Continuous inspection schemes. *Biometrika*, 41:100–115, 1954.

[18] A. Seyfried, B. Steffen, W. Klingsch, and M. Boltes. The fundamental diagram of pedestrian movement revisited. *Transportation Science*, 39(1):1–24, 2005.

[19] B. Steffen and A. Seyfried. Methods for measuring pedestrian density, flow, speed and direction with minimal scatter. *Physica A*, 389(9):1902–1910, 2010.

[20] K. Still. *Crowd Dynamics*. PhD thesis, Univ. of Warwick, 2000.

[21] D. Sugimura, K. Kitani, T. Okabe, Y. Sato, and A. Sugimoto. Using individuality to track individuals: Clustering individual trajectories in crowds using local appearance and frequency trait. In *ICCV*, 2009.

[22] B. Zhan, D. Monekosso, P. Remagnino, S. Velastin, and L. Xu. Crowd analysis: a survey. *Machine Visions and Applications*, 19(5–6):345–357, 2008.

[23] T. Zhao and R. Nevatia. Tracking multiple humans in crowded evnironment. In *CVPR*, 2004.