PFG 2012 / 6, 0701 – 0710
Stuttgart, December 2012

Article

Graph-Based Analysis of Pedestrian Interactions and Events Using Hidden Markov Models

Florian Burkert & Richard Bamler, München

Keywords: event detection, pedestrian surveillance, aerial image sequences, hidden Markov models, UAV

Summary: In this paper, we present an improved approach for the analysis of pedestrian interaction in crowded and cluttered scenes from aerial image sequences. Related work is limited to the detection of an undeclared abnormal event with regard to the common behaviour or to the detection of specific simple events incorporating only up to two trajectories. In our approach, event detection in pedestrian groups is done by detecting universal motion interaction patterns between pairs of pedestrians in a graph-based framework. Event detection is performed by analyzing the temporal behaviour of the motion interaction, which is represented by edges in the graph, by means of hidden Markov models (HMM). Temporarily disappearing edges in the graph can be compensated by an HMM buffer which internally continues the HMM analysis even if the corresponding pedestrians depart from each other awhile. Experimental results show the potential of our graph-based approach for event detection. The used datasets contain UAV image sequences in which an instructed pedestrian group simulates meaningful group behaviour and an aerial image sequence in which pedestrians approach a soccer stadium.

Zusammenfassung: *Graphenbasierte Ereignisdetektion von Fußgängerinteraktion mittels Hidden Markov Modellen*. In diesem Beitrag wird eine verbesserte Methode für die Detektion von Fußgänger-Interaktion in dichten und unstrukturierten Szenen aus Luftbildsequenzen vorgestellt. Bislang bestehende Arbeiten beschränken sich auf die Erkennung von unnormalen Ereignissen im Allgemeinen oder auf die Erkennung von einfachen Ereignissen, welche nur von bis zu zwei Personen durchgeführt werden. In der hier vorgestellten Methode wird Ereignisdetektion in Personengruppen vollzogen, wofür die Bewegungsinteraktion zwischen benachbarten Personenpaaren in einem graphenbasierenden System analysiert wird. Das zeitliche Verhalten der Bewegungsinteraktion wird mittels Hidden Markov Modellen (HMM) ausgewertet. Zeitlich unbeständige Kanten im Graph werden mit Hilfe eines HMM-Puffers abgefangen, welcher die Auswertung intern weiterführt, wenn sich das einer Kante zugehörige Personenpaar kurzzeitig voneinander entfernt. Es werden Ergebnisse präsentiert, welche das Potential der vorgestellten Methode zur Ereignisdetektion aufzeigen. Die verwendeten Datensätze beinhalten UAV-Sequenzen, welche Gruppenbewegungen eingewiesener Testpersonen beinhalten, und Luftbildsequenzen, welche Fußgänger vor einem Fußballstadion zeigen.

1 Introduction

The main objective of this work is event detection in crowds by robustly analyzing a pedestrian interaction graph using Hidden Markov Models in which the edges represent motion interaction patterns between pedestrians.

The huge amount of surveillance data requires automatic or at least semi-automatic interpretation. Consequently, research in crowd analysis has been intensified in the last decades in order to support human surveillance operators. In addition to purely image based crowd analysis techniques, crowd models from psychology, physics or from the nature have to be incorporated into more sophisticated surveillance systems (Zhan et al. 2008, Butenuth et al. 2011). Aerial imagery pro-

www.schweizerbart.de
DOI: 10.1127/1432-8364/2012/0150
1432-8364/12/0150 $ 2.50

vides a wide overview over a scene and can, therefore, ideally be used to extract trajectories of pedestrians which can then be used for event detection.

Numerous publications indicate the importance of crowd analysis. Zhan et al. (2008) present a survey which recapitulates contributions to object detection, tracking and event detection. The main input data for event detection are either trajectories or optical flow. Event detection systems using optical flow are able to detect abnormal events in high density crowds after observing the common behaviour for some time (Adam et al. 2008, Andrade et al. 2006, Mehran et al. 2009). However, individual behaviour cannot be inferred by optical flow and no classification of the type of the unusual event is made besides of flow-specific characteristics. In scattered and medium-dense scenes, the analysis of discrete trajectories is preferred because of the visibility of individuals. For the analysis of discrete trajectories, hidden Markov models (HMM) (Rabiner 1989) have often been applied in the past, which have originally been developed for speech recognition. Several specialisations of HMM have been developed and utilized for event detection, such as coupled HMM (CHMM) (Oliver et al. 2000) or switched dynamical HMM (SD-HMM) (Nascimento et al. 2010). Nascimento et al. (2010) classify recurring human trajectories in busy scenes by concatenating a given set of low level models using switched dynamical hidden Markov models (SD-HMM). Human trajectory mining is performed in the work of Calderara & Cucchiara (2010) by clustering frequent behaviours of pedestrians using different similarity measures. Kuettel et al. (2010) automatically learn spatio-temporal dependencies of moving agents and show experimental results from traffic scenes. However, by classifying or mining recurring trajectories, only a very stringent model containing some trajectory clusters can be built which is not flexible enough to cope with individual and spontaneous motion patterns in cluttered scenes that do not match recurring paths. Learning of recurring trajectories can also be used for the detection of unusual events (Basharat et al. 2008, Hu et al. 2006, Porikli & Haga 2004). Unusual events can only be detected if enough data about usual events is available beforehand. To this end, the scene has to provide specific conditions which can be followed by the majority of the observed objects, like entrance doors for pedestrians or driving lanes for vehicles.

We overcome the limitations of the related work. We use manually generated trajectories in order to be able to draw significant information about individuals' motion behaviour. The analysis of the entire scene is achieved by modelling all pedestrians in a graph at each frame. We calculate four extended motion features adapted from Burkert & Butenuth (2011) to deduce six universal motion patterns for each pair of trajectories. The motion patterns which describe the motion interaction of a pair of pedestrians correspond to the edges in the graph. The sequential behaviour of the motion patterns is analyzed by HMM. We focus on the detection of simple and universal motion patterns which allows us to interpret resulting large scale clusters of motion patterns but also individual events in the scene. At this level, findings from social and traffic sciences such as the social force model (SFM) Helbing & Molnár (1995) can be used. We show experimental results for our event detection system based on a dataset acquired by an unmanned aerial vehicle (UAV) in which an instructed group of pedestrians fulfills meaningful scenarios of group behaviour. Another real-world dataset contains an airborne image sequence in which pedestrians approach a soccer stadium.

The outline of this paper is as follows: In section 2 we introduce the terminology of HMM we use in this paper. Section 3 describes our system for robust, graph-based event detection. In section 4 we show experimental results and in section 5 we conclude and discuss future work.

2 Terminology of Hidden Markov Models (HMM)

A hidden Markov model (HMM) is a probabilistic model which is represented by a directed acyclic graph. A HMM shows the simplest form of a dynamic Bayesian network. The system underlying the HMM is a Markov chain of hidden states. At each time step, an observable output of the model is generated which only depends on the current hidden state.

2.1 *Parameters of an HMM*

An HMM provides clear Bayesian semantics and is defined by the following set of parameters (Rabiner 1989):

- A set of N possible hidden states $\{s_1, s_2, \ldots, s_N\}$, the state at time t is denoted as q_t.
- A set of M possible observations $\{v_1, v_2, \ldots, v_M\}$, the observation at time t is denoted as o_t.
- The transition probability matrix A with its elements a_{ij} denoting the transition probabilities from s_i to s_j

$$a_{ij} = P(q_t = s_j \,|\, q_{t-1} = s_i) \tag{1}$$

$$\sum a_{ij} = 1 \forall i, \quad a_{ij} \geq 0 \tag{2}$$

- The observation probability distribution $b_j(v_k)$ for an observation v_k

$$b_j(v_k) = P(o_t = v_k \,|\, q_t = s_j) \tag{3}$$

$$\sum_k b_j(v_k) = 1 \; \forall i, \quad b_j(k) \geq 0 \tag{4}$$

- The initial probabilities π_i that s_i is the initial state

$$\pi_i = P(q_1 = s_i) \tag{5}$$

$$\sum_i \pi_i = 1 , \quad \pi_i \geq 0 \tag{6}$$

From this set of parameters the transition probability matrix A, the observation probabilities $b_j(v_k)$ and the initial probabilities π_i can be subsumed under a parameter λ which characterizes an HMM.

2.2 *Inference in HMM*

The inference in HMM to find the most probable sequence of hidden states $\{q_1,q_2,\ldots,q_T\}$ is performed by using the corresponding given sequence of observations $\{o_1,o_2,\ldots,o_T\}$, where T is the length of the sequence. This problem can be solved by the *Viterbi algorithm* (Rabiner 1989) which is used if a complete and terminated sequence of observations is available. For our problem of event detection in real-time, which operates as the image sequence is acquired, filtering has to be used instead of the Viterbi algorithm. Filtering is used for computing the probability distribution over the hidden states $\{s_1,s_2,\ldots,s_N\}$ at a certain time step t, given the sequence of observations up to this time step $\{o_1,o_2,\ldots,o_{t-1},o_t\}$. Filtering can efficiently be solved by the *forward algorithm* (Rabiner 1989). The forward algorithm is appropriate for our task because it does not depend on an already terminated sequence and, thus, can be iteratively applied at every new frame of the image sequence.

The forward algorithm employs forward variables $\alpha_t(s_i), 1 \leq i \leq N$ which are calculated at each time step t for every hidden state s_i. Thus, the forward variables are defined as

$$\alpha_t(s_i) = P(o_1, o_2, \ldots, o_t, q_t = s_i \,|\, \lambda) \; . \tag{7}$$

$\alpha_t(s_i)$ is the probability to produce the observation sequence up to t and to reach state s_i at time t, given the HMM λ. At the first time step $t = 1$, the initialization of the forward algorithm is realized by

$$\alpha_1(s_i) = \pi_i b_i(o_1) \; . \tag{8}$$

The initialization of the forward variables depends on the initial probabilities π_i and the first observation o_1. At further time steps t, $2 \leq t \leq T$, the recursion of the forward algorithm is performed by

$$\alpha_t(s_j) = b_j(o_t) \sum_i \alpha_{t-1}(s_i) a_{ij} \; . \tag{9}$$

The recursion step depends on the observation o_t and on all forward variables of the previous time step $\alpha_{t-1}(s_i)$, multiplied by their probabilities of transition to state s_j.

3 Graph-Based Event Detection

In order to analyze motion interaction patterns in crowds we create a pedestrian interaction graph which contains all pedestrians of a scene. The analysis is done for existing edges in the graph by HMM-based event detection which is robust to fluctuant and partially departing pedestrians.

3.1 *Pedestrian Motion Model for Pedestrian Interaction*

The motion model consists of four motion features which are refined and extended versions of those defined in BURKERT & BUTENUTH (2011). The motion features are the observations for the HMM because they can ideally be used to infer six universal motion patterns. The motion patterns are treated as the events which have to be detected and are then used to interpret a crowd's behaviour.

Motion Features

We use four motion features which are calculated from trajectory pairs at every time step. A pair of trajectories is defined by two pedestrians i and j which are sufficiently close to each other such that a significant motion interaction takes place. The method to specify significant motion interaction is described in section 3.2. The four motion features for two pedestrians i and j are the sum of the velocities v_i+v_j, the variation of the distance Δd, the average pedestrian density $\mu(\rho_i+\rho_j)$ and the normalized scalar product of both motion direction vectors s. The motion features serve as the observations in the HMM. Fig. 1 (left) depicts two trajectories illustrating the motion features v_i+v_j, Δd and $\mu(\rho_i+\rho_j)$. The velocity v_i of a pedestrian is calculated at each time step using the frame rate and the covered distance since the last time step. The variability of the distance is defined as $\Delta d=d_t / d_{t-1}$, with d_{t-1} being the distance at time step $t\text{-}1$ and d_t being the distance at time step t. Thus, $\Delta d > 1$ for an increasing distance and $\Delta d < 1$ for a decreasing distance. The local pedestrian density ρ_i is calculated by the inverse of the area of the corresponding cell in a Voronoi diagram, which is constructed from the pedestrian locations at each frame (STEFFEN & SEYFRIED 2010). By using a Voronoi diagram, the local pedestrian density can be calculated instead of using a fixed area in which the number of pedestrians is counted. Only for those pedestrians which are located at the border of groups, the density is calculated by counting the number of pedestrians within a specified radius r and relating it to the area. The reason for this exception is that Voronoi cells at the border of a point cluster will receive very large or infinite area (Fig.1, right). We further introduce a fourth motion feature which is the normalized scalar product s of both motion direction vectors. s receives values up to 1 for pedestrians walking in parallel (Fig. 1, left), values of about 0 for orthogonal vectors, and up to -1 for pedestrians walking in opposite directions. Thus, s complements the feature Δd by describing the type of the distance variation. Δd only gives the absolute variation of the distance between two pedestrians but does not specify the directions, in which the distance variation is fulfilled. In summary, the feature vector is $[v_i+v_j, \Delta d, \mu(\rho_i+\rho_j), s]$.

Motion Patterns

We define six motion patterns which occur when pedestrians are close to each other, namely *Standing*, *Queueing*, *Walking*, *Running*, *Diverging* and *Converging*. These simple and universal motion patterns are the basis for our event detection system and define the type of the motion interaction between neighbouring pedestrians. The speed of pedestrians

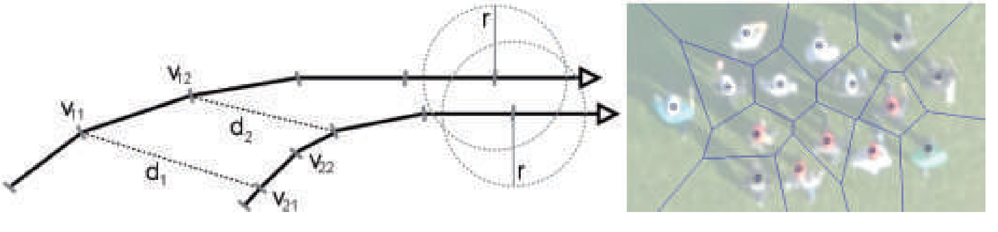

Fig. 1: Motion features derived from trajectories; left: v_{it} is the velocity of trajectory i at time t, d_t is the distance at frame t and r is the radius in which the pedestrian density at a group boundary is computed; right: Voronoi diagram of pedestrian locations, pedestrian density is defined as the inverse cell size.

can be variable at the motion patterns *Converging* and *Diverging* and is at most 0.1, 0.3, 2.0 and 4.5 m/s for the motion patterns *Standing*, *Queueing*, *Walking*, *Running*, respectively. The variability of distance explicitly leads to a statement if two pedestrians approach or depart from each other, independent of the motion direction. Therefore, the variability of distance is about 1 for *Standing* and the parallel motion patterns *Queueing*, *Walking* and *Running*. Consequently, $\Delta d > 1$ for *Diverging* and $\Delta d < 1$ for *Converging*. The pedestrian density is variable but rather low when pedestrians are standing or moving naturally and can reach high values up to 5 pedestrians per m^2 in dense queuing areas. The normalized scalar product of the motion direction vectors is variable for standing pedestrians because a slight motion into a random direction always occurs and no pedestrian will stand completely still. The motion patterns *Queueing*, *Walking* and *Running* are characterized by pedestrians walking in parallel, so the scalar product is close to 1. The scalar product emphasizes parallel scenarios in which the variability of distance might misleadingly suggest diverging or converging motion patterns because of different velocities.

Usually, training data from the real world surveillance scenes is used to learn HMM offline. However, we do not use data from surveillance cameras to learn the HMM for event detection because we focus on cluttered scenes which occur at big events. Learning from real world data always relies on frequently recurring motion paths within the scene of interest, which we cannot assume to be available for any place where a big event may take place. Instead, we use synthetic data representing the motion patterns to learn the HMM. The training data are generated by moving agents which follow rules of motion depending on the description above. The training data consist of 900–1200 feature vectors per motion pattern which were calculated from normally distributed simulated trajectories. From these feature vectors, the mean values are used to derive the feature vector of each motion pattern. Based on the central limit theorem, the feature vectors are approximately normally distributed. For the estimation of the hidden state q_t which corresponds to the motion pattern, a grid-based ML estimation is used. The initial probabilities π_i are assumed to be uniformly distributed. The transition probabilities a_{ij} are set manually in a way to reflect the fact that human motion is very variable, so there is no regular scheme if a pedestrian might stop, turn left or turn right after walking straight. The values used for the transition probabilities a_{ij} are thus nearly identical, with deviations from a uniform distribution in the range of 0.05 to 0.10. For example, these deviations model that after *Standing*, *Running* is less probable than *Walking*. By setting the transition probabilities a_{ij} in the way just described, we overcome the limitation that there exists no real world or synthethic training data which represents realistic transition probabilities between the motion patterns in our approach.

3.2 *Analysis of a Pedestrian Interaction Graph*

Event detection in a scene is performed based on a pedestrian interaction graph in which nodes represent pedestrians and edges represent motion interaction between pairs of pedestrians. The motion interaction is modelled as pairwise motion patterns which are analyzed using HMM. The graph is capable of robustly changing its topology because it is dynamically updated at every frame.

Edge Weight for Graph Simplification

The number of edges and the computational cost for the analysis of the graph is $N \cdot (N\text{-}1)/2$ for a number of N pedestrians. To overcome this high computational cost, we introduce edge weights based on a Gaussian function including the pedestrian density to significantly reduce the number of edges in the graph. Thus, only edges representing significant motion interactions between directly adjacent pedestrians are considered in the graph. The weight function $w_{ij}(d,\rho)$ with $0 \leq w_{ij} \leq 1$ between two nodes i and j is defined as

$$w_{ij} = \exp\left(-\frac{d_{ij}^{\,2}}{\frac{1}{2\rho}} \right). \tag{10}$$

The weight w_{ij} depends on the distance d_{ij} between the pedestrians representing the nodes i and j and on the density ρ which is given by pedestrians per m^2. The weight function is a Gaussian function with height 1, μ=0 and $\sigma = 1/\left(\sqrt{\rho} \cdot 2\right)$. Thus, the weight w_{ij} receives high values for directely adjacent pedestrians i and j where a significant motion interaction is supposed to take place. At high pedestrian densities the weight w_{ij} is only high between pedestrians that have a distance of a few decimeters, whereas at low densities the weight w_{ij} can be high even if adjacent pedestrians have an offset of several meters. We introduce a threshold which is applied to the weights in order to determine which edges in the graph are to be constructed and which edges are omitted.

Framewise Graph Updating

Our graph-based approach for event detection in crowds represents dynamic behaviour of pedestrians. To this end, the pedestrian interaction graph is capable of changing its topology at every frame depending on the new arrangement of pedestrians in the scene. Fig. 2 shows an example of four pedestrians and their trajectories. Particular time steps are represented by dotted lines in between the trajectories. Figs. 2 a), 2 b) and 2 c) show the corresponding graph with four nodes. The topology changes by inserting a new edge between nodes 2 and 3 because of the decreased distance between the corresponding pedestrians. The density is supposed to be constant. The width of the edge connecting nodes 2 and 3 increases in Fig. 2 c) as a consequence of the increased weight w_{23}.

There are three preconditions of how our system deals with the sequential interaction analysis, depending on the configuration of the graph in the previous step. The first case is that an edge existing in the previous time step will further exist, such that the corresponding interaction analysis can be continued. The second case is that two pedestrians are converging and the weight w_{ij} exceeds the threshold. In this case, a new edge is generated and the analysis of this interaction is started. The third case is that two pedestrians diverge and the weight w_{ij} falls below the threshold. In this case, the corresponding edge is removed from the pedestrian interaction graph.

Robust Event Detection

The interaction analysis between a pair of pedestrians is performed by HMM. The forward algorithm is used to derive the type of motion pattern for each pair of pedestrians, for which a common edge in the graph exists. When applying the forward algorithm, the motion features serve as observations and the motion patterns serve as the hidden states of the HMM. Edges in the graph can arise or disappear during the sequence because of the dynamic behaviour of the crowd described in the previous section. Pedestrians do not move in a linear way but tend to slightly deviate to the left or right while walking. Therefore, the interaction analysis bears the risk of being interrupted for some frames if pedestrians depart from each other for a short time only. To overcome this risk and to achieve a robust sequential analysis of the motion interaction, an HMM buffer is used when analyzing the pedestrian interaction graph. The HMM buffer is internally activated for a specific interaction when the weight of the corresponding edge decreases below the threshold. At this point, the recursion of the forward algorithm would be terminated if no HMM buffer was used. However, the recursion is internally continued for a user-defined maximal number of frames in the HMM buffer. In the case that the two cor-

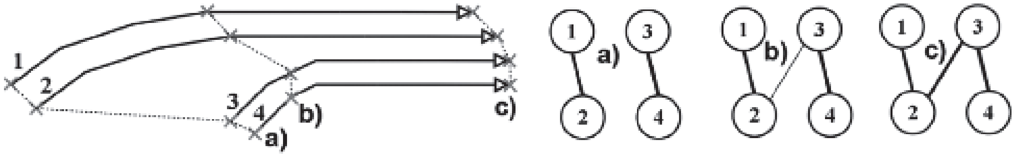

Fig. 2: Left: graph updating for four synthetic trajectories 1–4; right: three representative graphs showing the topology at particular frames (a, b, c) related to the dotted lines.

responding pedestrians approach each other and the weight increases again, the consistently and eternally analysed motion interaction is loaded from the HMM buffer and the result is subsequently added to the corresponding interaction analysis. Hence, the temporarily omitted corresponding edge of the graph is subsequently constructed. Thus, no fragments of the corresponding pedestrian interaction can arise. If the weight does not increase again after the defined number of frames, the corresponding interaction analysis is terminated and the edge is finally deleted.

4 Experimental Results

We present experimental results for robust pedestrian interaction analysis using two datasets with different camera platforms and different scenes. The datasets are described in the next section. Afterwards the results show significant scenes and demonstrate the robustness of our approach.

4.1 *Datasets*

The first dataset used contains image sequences captured from a UAV octocopter platform. The images were taken from a flying height of 85 m with a Panasonic DMC-LX3 camera, resulting in a ground resolution of 1.5 cm. The frame rate of the image sequences is 1 Hz. The captured scenes contain more than 10 different scenarios of pedestrian group behaviour in different complexity levels. The pedestrians were instructed about the scenarios in advance; however, the information was reduced to a minimum in order to preserve natural behaviour. The captured scenes contain simple scenarios such as commonly walking pedestrians but also complex scenes like a bottleneck, crossing pedestrian groups at different velocities or an escaping situation. The second dataset is an image sequence taken by an airborne camera platform of the German Aerospace Center (DLR). The image sequence contains 16 frames taken at a frame rate of 2 Hz. The ground resolution is 0.15 m at a flying height of 1000 m. The area of interest is a soccer stadium where thousands of people are approaching the gates.

For the experimental results we use pedestrian trajectories which were generated manually from the image sequences because we focus on realistic trajectories and the potential of our graph-based event detection system for realistic pedestrian behaviour. However, our event detection system is able to deal with possibly incomplete automatically generated tracklets because the pedestrian interaction graph can deal with changing topology in a straightforward way.

4.2 *Event Detection Results*

The event detection results for the UAV dataset are shown in Figs. 3 and 4, including a colourbar which depicts the colour labels for the edges corresponding to the motion patterns. In Fig. 3, a group of pedestrians passes a narrow bottleneck (frames 3, 6, 9 and 12 are shown). Our event detection system successfully detects the typical motion interaction characteristics. Neighbouring pedestrians *Converge* and *Walk* towards the bottleneck, which is illustrated by orange and blue edges in frames 3 and 6. Pedestrians who have passed the gap *Diverge* and *Walk* ahead, whereas the pedestrians at the back of the group have to *Queue* for a while in frame 9. In Fig. 4, a corridor scenario of two walking groups walking in opposite directions is successfully detected. This scenario is characterized by two approaching and internally *Converging* groups in which the backmost pedestrians again have to *Queue* because of the narrowness of the corridor. The formation of lanes, which has already been investigated in Helbing et al. (2001), can be confirmed by the linearly arranged motion pattern *Walking* and the oppositely arranged motion patterns *Converging* and *Diverging* (frames 8 and 11).

The results for the soccer stadium sequence are presented in Figs. 5 and 6. Our event detection system is robust in the case that interacting pedestrians depart from each other for a short time by applying the HMM buffer. This robustness is exemplified in Fig. 5: the top row contains three consecutive frames of a pedestrian interaction graph where edges between node 7 and nodes 5 and 6 are not present in the middle frame. This graph was produced without the HMM buffer. The bottom row shows the graph which was produced with the HMM buffer. Here, the corresponding edges exist such that a continuous and robust analysis of the interaction between the pedestrians 5, 6 and 7 can be performed. In Fig. 6, *Queueing* pedestrians are successfully detected, which is displayed by the yellow edges during the whole sequence. Some pedestrians are passing the queue and perform multiple interactions due to freedom of motion. During the sequence, the density in the narrow area between

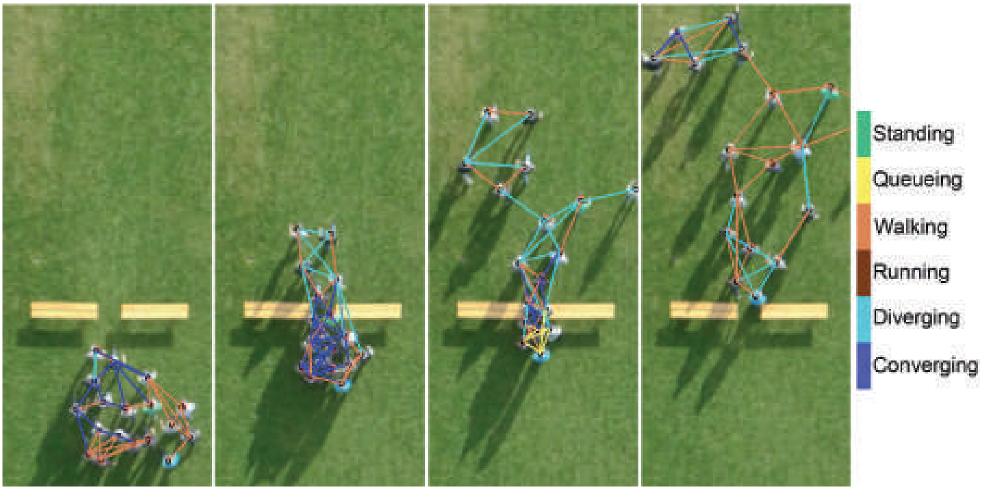

Fig. 3: Event detection result (frames 3, 6, 9 and 12) for the UAV dataset showing a bottleneck scenario in the upwards walking pedestrian group.

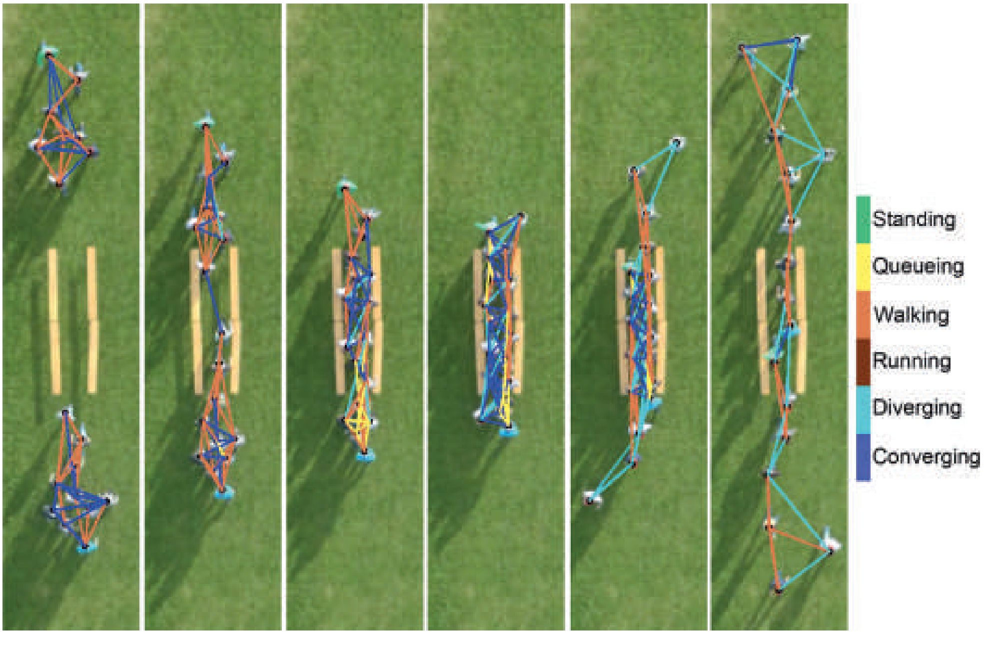

Fig. 4: Event detection result (frames 2, 5, 8, 11 and 14) for the UAV dataset showing a corridor scenario between two antiparallel walking pedestrian groups.

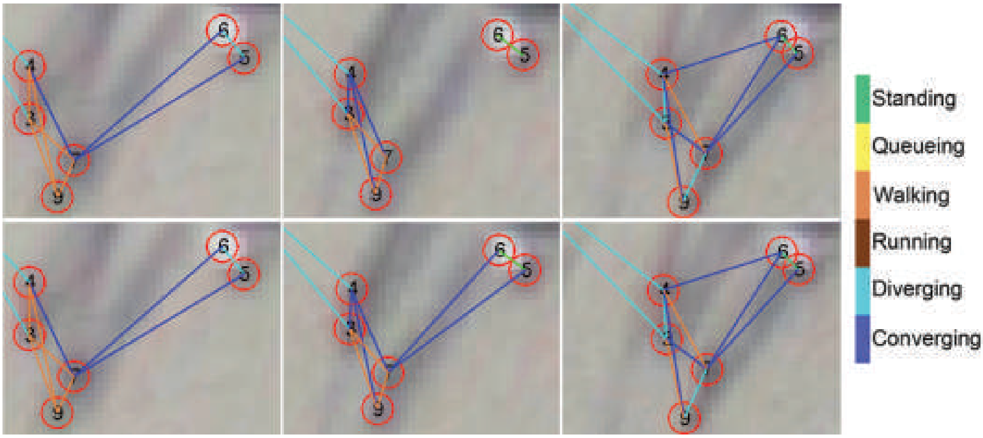

Fig. 5: Top: pedestrian interaction graph without HMM buffer; bottom: graph with HMM buffer.

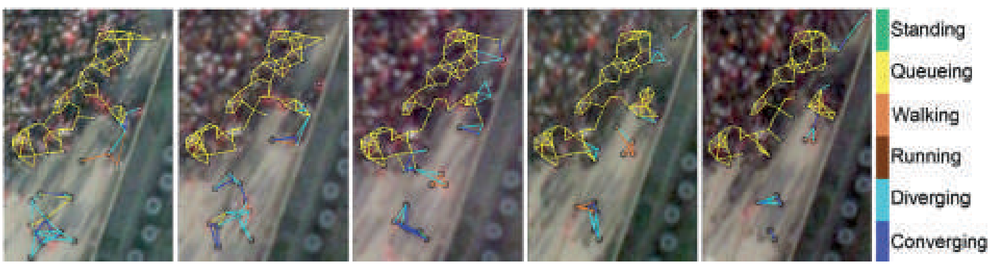

Fig. 6: Event detection result (frames 2, 6, 10, 14 and 16) for the soccer stadium dataset showing pedestrians passing a queue in a narrow area by a wall.

the queue and the wall on the right rises and the velocity decreases, which is illustrated by more and more *Queueing* patterns in this area.

The results demonstrate the potential of our system for graph based event detection by analyzing motion interaction in groups of pedestrians. Using HMM, the sequential behaviour of motion interaction between two pedestrians can reliably be analyzed. The six motion patterns we have defined represent human motion behaviour in a simple manner, such that areas of homogeneous behaviour as well as specific behaviour of only a few pedestrians can be inferred. In some frames outliers of the predominant motion pattern for a continuous pedestrian interaction arise. This is caused by the individual freedom of motion of pedestrians. Single and momentary variations of motion interaction in crowded scenes have to be expected. The low frame rate of the UAV dataset is sufficient to derive motion features which are used for event detection.

5 Conclusions

In this paper, we present a new approach for the analysis of pedestrian interaction and events in crowded scenes. We construct a pedestrian interaction graph in which nodes represent all pedestrians in a scene and edges represent motion interaction between neighbouring pedestrians. The graph can change its topology during the sequence and is robust against fluctuating and partly departing pedestrians. A set of six universal motion patterns is defined, describing the type of interaction which is then detected by HMM. We extend the motion features of previous work by the scalar product of motion directions and a refined density calculation, serving as observations for the HMM. We use a new UAV dataset for the evaluation of our event detection system, as well as an aerial dataset of a soccer stadium. The promising results show the potential of our approach to interpret pedestrian motion behaviour. Future work aims at a higher level analysis of the graph for an automatic and probabilistic detection of complex events in pedestrian groups.

Acknowledgements

We thank the German Aerospace Center (DLR) for providing the image sequence and for funding this work and also the International Graduate School of Science and Engineering (IGSSE) of Technische Universität München for supporting this study.

References

Adam, A., Rivlin, E., Shimshoni, I. & Reinitz, D., 2008: Robust Real-Time Unusual Event Detection using Multiple Fixed-Location Monitors. – IEEE Transactions on Pattern Analysis and Machine Intelligence **30** (3): 555–560.

Andrade, E.L., Blunsden, S. & Fisher, R.B., 2006: Hidden Markov Models for Optical Flow Analysis in Crowds. – **18th** International Conference of Pattern Recognition, Hong Kong, 460–463.

Basharat, A., Gritai, A. & Shah, M., 2008: Learning Object Motion Patterns for Anomaly Detection and Improved Object Detection. – IEEE Conference on Computer Vision and Pattern Recognition: 1–8, Anchorage, USA.

Burkert, F. & Butenuth, M., 2011: Event Detection based on a Pedestrian Interaction Graph using Hidden Markov Models. – Stilla, U., Rottensteiner, F., Mayer, H., Jutzi, B. & Butenuth, M. (eds.): Photogrammetric Image Analysis. – Lecture Notes in Computer Science **6952:** 271–283, Springer, Berlin, Germany.

Butenuth, M., Burkert, F., Borrmann, A., Kneidl, A., Hinz, S., Schmidt, F., Sirmacek, B. & Hartmann, D., 2011: Integrating Pedestrian Simulation, Tracking and Event Detection for Crowd Analysis. – International Conference on Computer Vision Workshops, MSVLC, Barcelona, Spain.

Calderara, S. & Cucchiara, R., 2010: People Trajectory Mining with Statistical Pattern Recognition – **23rd** IEEE Conference on Computer Vision and Pattern Recognition Workshops, San Francisco, USA.

Helbing, D. & Molnár, P., 1995: Social Force Model for Pedestrian Dynamics. – Physical Review E **51** (5): 4282–4286.

Helbing, D., Molnár, P., Farkas, I.J. & Bolay, K., 2001: Self-Organizing Pedestrian Movement. – Environment and Planning B: Planning and Design **26:** 361–383.

Hu, W., Xiao, X., Fu, Z., Xie, D., Tan, T. & Maybank, S., 2006: A system for learning statistical motion patterns. – IEEE Transactions on Pattern Analysis and Machine Intelligence **28** (9): 1450–1464.

Kuettel, D., Breitenstein, M.D., Van Gool, L. & Ferrari, V., 2010: Learning Object Motion Patterns for Anomaly Detection and Improved Object Detection. – **23rd** IEEE Conference on Computer Vision and Pattern Recognition, San Francisco, USA.

Mehran, R., Oyama, A. & Shah, M., 2009: Abnormal Crowd Behaviour using Social Force Model. – **22nd** IEEE Conference on Computer Vision and Pattern Recognition, Miami Beach, USA.

Nascimento, J.C., Figueiredo, M.A.T. & Marques, J.S., 2010: Trajectory Classification using Switched Dynamical Hidden Markov Models. – IEEE Transactions on Image Processing **19** (5): 1338–1348.

Oliver, N.M., Rosario, B. & Pentland, A.P., 2000: A Bayesian Computer Vision System for Modeling Human Interactions. – IEEE Transactions on Pattern Analysis and Machine Intelligence **22** (8): 831–843.

Porikli, F. & Haga, T., 2004: Event Detection by Eigenvector Decomposition using Object and Frame Features. – **17th** IEEE Conference on Computer Vision and Pattern Recognition Workshops, Washington, DC, USA.

Rabiner, L.R., 1989: A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition. – IEEE **77** (2): 257–286.

Steffen, B. & Seyfried, A., 2010: Methods for measuring pedestrian density, flow, speed and direction with minimal scatter. – Physica A: Statistical Mechanics and its Applications **389** (9): 1902–1910.

Zhan, B., Monekesso, D.N., Remagnino, P., Velastin, S.A. & Xu, L., 2008: Crowd Analysis: A Survey. – Machine Vision and Applications **19** (5–6): 345–357.

Address of the Authors:

Dipl.-Ing. Florian Burkert and Prof. Dr. Richard Bamler, Technische Universität München, Lehrstuhl für Methodik der Fernerkundung, D-80333 München, Tel.: +49-89-289-22643, Fax: +49-89-289-23202, e-mail: florian.burkert@bv.tum.de, richard.bamler@dlr.de

Manuskript eingereicht: März 2012
Angenommen: August 2012