

COMPLEX EVENT DETECTION IN PEDESTRIAN GROUPS FROM UAVS

F. Burkert *, M. Butenuth

Technische Universität München, Remote Sensing Technology, 80333 München, Germany
(florian.burkert, matthias.butenuth)@bv.tum.de

Commission III, WG III/5

KEY WORDS: Event detection, pedestrian groups, aerial image sequences, UAV

ABSTRACT:

We present a new hierarchical event detection approach for highly complex scenarios in pedestrian groups on the basis of airborne image sequences from UAVs. Related work on event detection for pedestrians is capable of learning and analyzing recurring motion paths to detect abnormal paths and of analyzing the type of motion interaction between pairs of pedestrians. However, these approaches can only describe basic motion and fail at the analysis of pedestrian groups with complex behavior. We overcome the limitations of the related work by using a dynamic pedestrian graph of a scene which contains basic pairwise pedestrian motion interaction labels in the first layer. In the second layer, pedestrian groups are analyzed based on the dynamic pedestrian graph in order to get higher-level information about group behavior. This is done by a heuristic assignment of predefined scenarios out of a model library to the data. The assignment is based on the motion interaction labels, on dynamic group motion parameters and on a set of subgraph features. Experimental results are shown based on a new UAV dataset which contains group motion of different complexity levels.

1. INTRODUCTION

The main objective of this work is hierarchical event detection in pedestrian groups based on UAV image sequences in order to recognize complex scenarios automatically.

Crowd monitoring and analysis are important research fields because the huge amount of surveillance data requires automatic or at least semi-automatic interpretation. As a corollary of the need for automatic surveillance systems, image-based crowd analysis techniques and crowd models from physics or nature have to be combined in order to achieve more intelligent surveillance systems (Zhan et al., 2008; Butenuth et al., 2011). Furthermore, it is essential to incorporate findings from psychological or social crowd studies into surveillance systems. Pedestrians behave individually and not like computable robots, however, pedestrian group behavior follows specific self-organizing rules. For example, the formation of lanes with homogeneous motion direction in oppositely moving groups can be observed (Helbing et al., 2001). For individuals, measurable distance definitions for people interaction as well as the correlation between physical and social distances have been proposed in the proxemics theory (Hall, 1966). The interaction of individuals with groups has been analyzed in studies about group size and social boundaries (Mullen et al., 1991).

Past work in the field of image-based event detection is focused on the detection of abnormal events of single pedestrians or in a moving crowd. Pedestrian trajectories are a common subject for classification and analysis or the basis for event detection. Oliver et al. (2000) model human motion interaction by analyzing a set of secondary output features of two pedestrian trajectories (such as the orientation and the velocity) using coupled Hidden Markov Models (CHMM). Despite of being capable of recognizing complex pairwise events, this approach

fails for large groups of pedestrians. 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). However, many datasets containing recurring trajectories are required to apply this approach, as for example at shopping malls or parking lots. In both approaches, the sequential characteristics of trajectories are analyzed with specific extensions of Hidden Markov Models (HMM) (Rabiner, 1989). Trajectory learning can also be used for the detection of outliers from the expected typical motion patterns (Hu et al., 2006; Porikli and Haga, 2004). An unusual event is marked as not belonging to the common motion patterns, but no specific classification or meaning is assigned to it. Additionally, trajectory classification or learning cannot be used for complex event detection in groups of pedestrians which move individually and without uniquely recurring behavior. In contrast, optical flow has been used for unusual event detection in groups of pedestrians instead of pedestrian trajectories (Mehran et al., 2009). This approach is able to detect unusual events when monitoring a moving crowd after some time of typical motion. However, no influence of unusual motion of a single pedestrian is discriminable and no classification of the unusual event is made besides of flow-specific characteristics.

In this paper we overcome the limitations of the related work in order to perform complex event detection in groups of pedestrians hierarchically. We use trajectories of single pedestrians as input to be able to consider individual motion and, thus, the resulting specific motion interactions between neighboring pedestrians. In the first layer we construct a dynamic pedestrian interaction graph as proposed in Burkert et al. (2011), with refined and extended motion interaction features. In the second layer, we extract higher-level information about group behavior. We compare the graph

* Corresponding author. This is useful to know for communication with the appropriate person in cases with more than one author.

gained from the data with a scenario model library which consists of common group behavior up to very specific and dangerous group behavior as for example an “*escaping*” scenario in Figure 1. The complex event detection is done by a heuristic search based on connected components in the graph. Used group motion parameters are the motion interaction labels, group-related features (which are represented by histograms, like speed) as well as subgraph features like the number of nodes. We show experimental results based on a new UAV dataset which contains image sequences of more than 10 different scenarios with pedestrian groups under a controlled setting.

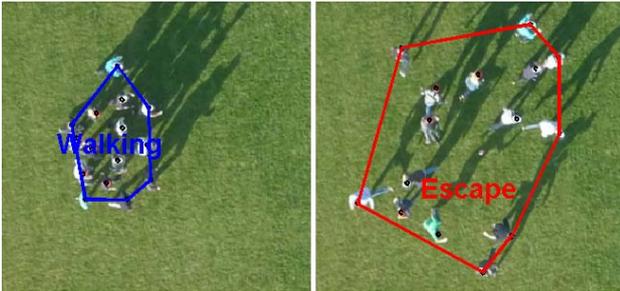


Figure 1. Example scenario “*escaping*” from the UAV dataset and the complex event detection result: normally walking group on the left, escaping group some frames later on the right.

The structure of this paper is as follows. In section 2 we describe the first layer and the construction of the pedestrian interaction graph. Section 3 contains our complex event detection layer, followed by experimental results in section 4. In section 5 we conclude and discuss our work.

2. PEDESTRIAN INTERACTION GRAPH

In the first layer of our hierarchical complex event detection approach, we construct a dynamic pedestrian interaction graph. Pedestrian trajectories are assumed to be available as input. The graph contains all pedestrians in the scene and can change its topology at each frame. It shows the type of the pairwise motion interaction between neighbouring pedestrians (Burkert et al. 2011). The model for pedestrian motion is described in the next subsection, followed by a compact overview over the graph-based pairwise motion analysis.

2.1. Motion Model

The motion interaction between neighboring pedestrians is analyzed by inferring the type of motion pattern between two neighboring trajectories, which itself is derived from a set of four pairwise motion features.

The first motion feature is the sum of the velocities of both pedestrians i and j , $\sum v_{ij}$. The second motion feature is the variation of the distance between both pedestrians $\Delta d = d_t / d_{t-1}$, with d_{t-1} being the distance at frame $t-1$ and d_t being the distance at frame t . Thus, $\Delta d > 1$ at an increasing distance and $\Delta d < 1$ at a decreasing distance. The third motion feature is the average pedestrian density around both pedestrians D_{ij} . We refine the method for calculating the local pedestrian density, which is typically given by $D = N / |A|$, where N is the number of pedestrians in an area

A. Instead of using a fixed area in which the number of neighbours is counted, we construct a Voronoi diagram from all pedestrians at each frame similar to Steffen et al. (2010). By doing so, we avoid getting only an average density in an area but a detailed local pedestrian density. Figure 2 shows an example for a Voronoi diagram in a group of 15 pedestrians. The size a_i of a Voronoi cell corresponds to the individual space of a pedestrian i . The inverse of the individual space gives the local pedestrian density which is given in pedestrian per square meter. Therefore, the third motion feature D_{ij} is calculated by

$$D_{ij} = \frac{1}{\frac{1}{2}(a_i + a_j)} \left[\frac{p}{m^2} \right]. \quad (1)$$

For Voronoi cells which are located at the border of a group and have infinite or unrealistic big size, we count neighboring pedestrians within a fixed area to calculate the density. 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 parallel walking pedestrians, values around 0 for orthogonal vectors, and up to -1 for antiparallel walking pedestrians (see Figure 3).

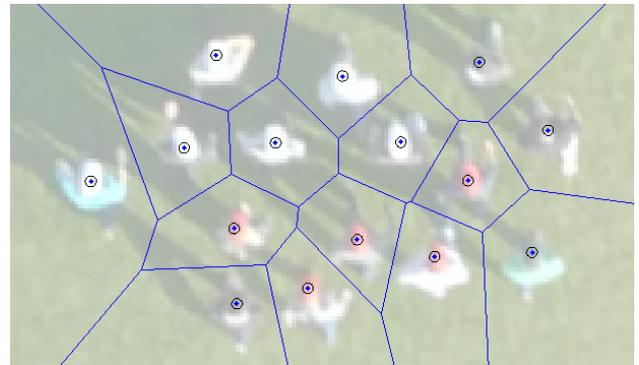


Figure 2. Voronoi diagram of a pedestrian group. Pedestrian density is defined as the inverse of the cell size. For infinite cell sizes at the boundary, an area-based density calculation is applied.

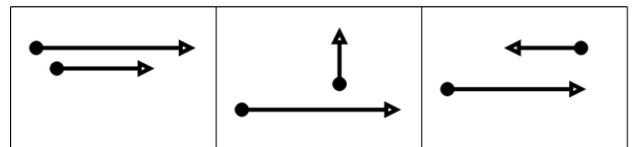


Figure 3. Normalized scalar product for moving pedestrians (fourth motion feature): $s = 1$ for parallel walking pedestrians (left), $s = 0$ for orthogonal walking pedestrians (center), $s = -1$ for antiparallel walking pedestrians (right).

We define six simple pairwise motion patterns which usually occur between neighboring pedestrians. Pairwise motion patterns are suitable for event detection in crowds, because they focus on motion interaction between pedestrians. In contrast, a single person walking on an open area has no motion interaction to other pedestrians and, thus, is of minor interest for event detection in groups. The six motion patterns are *together standing*, *together queuing*, *parallel walking*, *parallel running*, *diverging* and *converging*, each defined by a combination of specific intervals of the four motion features.

2.2. Graph-based pairwise motion analysis using HMM

The dynamic pedestrian graph is constructed for every frame of a sequence and contains all pedestrians in the scene. Nodes in the graph represent pedestrians and edges represent interactions between pedestrians. The dynamic pedestrian graph can change its topology at each frame and is flexible with regard to the number of included nodes. We only consider motion interaction between directly adjacent pedestrians by introducing a Gaussian weight function in which the width is depending on the local pedestrian density. Thus, only edges representing motion interaction between neighboring pedestrians are constructed.

Edges in the graph are labeled with the type of the current motion pattern between neighboring pedestrians which is inferred by HMM using the forward algorithm (Rabiner, 1989). When computing the forward algorithm, the motion features serve as observations and the motion patterns serve as the hidden states of the HMM. The detection of the current motion pattern is robust against temporally deviating neighboring pedestrians, a case that can occur frequently because of the individual motion options for pedestrians. In such cases, edges might be deleted for some frames. However, the dynamic pedestrian graph internally keeps a short history of the pertinent motion interaction and is able to rebuild the corresponding edge.

3. COMPLEX EVENT DETECTION IN GROUPS

In the second layer of our hierarchical approach we analyze the pedestrian interaction graph in combination with specific group-related motion features. Complex event detection is achieved by a heuristic assignment of predefined scenarios for group events from a model library to the data. Our approach for complex event detection is presented in the next section, followed by a description of the scenario model library.

3.1. Complex event detection by data-model comparison

Connected component analysis: The detection of complex events in pedestrian groups is based on the pedestrian interaction graph which was constructed in the first layer. The graph is incomplete, that is, not all edges exist and the graph may consist of several independent subgraphs. The main contribution of the pedestrian interaction graph to the second layer of our approach is to provide an indication of connected pedestrian groups and to deliver information about pairwise motion interaction. Therefore, connected components in the graph have to be detected in order to identify connected pedestrian groups in the scene. The detection of connected components is done in each frame by computing the Dulmage-Mendelsohn decomposition of the sparse adjacency matrix which represents the graph (Dulmage and Mendelsohn, 1958). The history of merging and splitting of connected components throughout the sequence is stored in a sparse history matrix in order to remember previous group motion. In this sparse history matrix, the rows denote the parent connected components and the columns denote the child connected components. Connected components consisting of only one pedestrian are labelled with index 1. An example for connected component analysis and history storage is depicted in Figure 4, where the connected component 2 splits into connected components 3 and 4 after the first frame. In the next frame, the connected component 3 is unchanged and the connected component 4 splits into the connected component 5 and the single node labeled with 1.

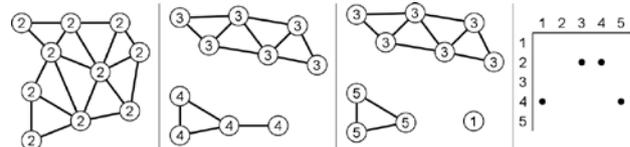


Figure 4. Example for connected component analysis: three subsequent graphs with connected component labels on the left, corresponding sparse history matrix on the right.

Parameterization of scenarios: The basis for our complex event detection approach is the pedestrian interaction graph in which connected components represent connected pedestrian groups. We define a set of group-related parameters which are derived from connected components but are flexible with regard to group size. These parameters are then used to identify specific group behaviour, that is, to assign scenarios from the model library to the data.

In more detail, the group motion parameters are:

- the normalized histogram of the six motion patterns within one connected component (described in Section 2.1.) which is derived directly from the pedestrian interaction graph.
- the number of nodes in each connected component: this feature is used in combination with the history matrix to decide whether one or more pedestrians join a group or separate from a group.
- the mean and standard deviation of the speed of all pedestrians in one connected component, given in m/s .
- the mean and standard deviation of the pedestrian densities within one connected component, given in p/m^2 .
- the normalized histogram of the scalar product s between motion direction vectors of neighboring pedestrians. The interval for s which ranges from -1 to 1 is divided up into 20 segments, each of it having width 0.1 (Section 2.1.).

A histogram normalization for the motion patterns and the scalar of the motion vectors is necessary because the complex event detection approach has to be independent of the number of pedestrians within a group. All parameters are defined such that scenarios are represented universally and the assignment of scenarios from the library is independent of the group size.

Scenario model assignment: Each of the scenarios in the library is uniquely defined by group motion parameters which are calculated based on connected components. The assignment of scenario models to the data is implemented by a complex heuristic search framework which tests previously calculated parameters of connected components for affiliation to scenarios from the model. The framework is hierarchical for most of the scenarios which can be assigned to the data in parallel or supplemental. Therefore, it is possible to assign more than one scenario in order to refine the information derived from the complex event detection approach. More information on parallel assignment of scenarios is given in section 3.2.

3.2. Scenario model library

We construct a model library of scenarios which may occur within moving pedestrian groups. This model library serves as database whose scenarios can be assigned to the data by the

complex event detection module. The construction of the model library is influenced by psychological and social crowd studies and by crowd modeling techniques in order to achieve realistic scenario models. On the one hand, scenarios in the library comprise simple and common group behavior which occurs in everyday life, like normal walking. On the other hand, scenarios in the library comprise complex events which are uncommon and possibly dangerous, like a bottleneck situation. Complex events are specified as group behavior that can occur when one or more groups of pedestrians interact with each other or with objects. In the following, the scenario model library is described in detail.

The most common scenarios are three basic group motion characteristics, namely *Standing*, *Walking* and *Running*, which are uniquely defined by the histogram of pairwise motion patterns and by constant speed- and density histograms. Furthermore, the scenarios *Standing* and *Walking* serve as supplemental information for many of the following scenarios, depending on the pace in which the specific scenario is carried out. The next common scenarios are *Accelerating* and *Slowing Down*, which are uniquely defined by the change of the speed histogram. More complex scenarios are *Individual Merge* or *Individual Split* and *Group Merge* or *Group Split*, depending on a change of connected components by one or more nodes in the graph. The split and merge scenarios are also possibly supplemental or co-existent with other scenarios because groups or individuals can separate or join continuously. A special case of a group merge or individual merge is the *Frontal Collision* scenario, in which pedestrians interfuse each other. In this scenario, the histogram of the scalar of the motion vectors uniquely notices a temporal rise of antiparallel moving pedestrians. Additionally, the histogram of pairwise motion patterns reflects the increase of converging interactions in contrast to the normally walking interaction shortly before. The formation of lanes at interfusing groups can be confirmed by still existing normally walking patterns between pedestrians of common motion direction and a large amount of converging and diverging interactions. The formation of lanes can be observed in narrow passages; therefore, this specific scenario is called *Corridor*. A possibly dangerous scenario is the *Bottleneck* scenario, which is defined by an interim increase in the density histogram, a decrease in the speed histogram and change from converging to diverging in the histogram of pairwise motion patterns. Also, an *Escape* of a walking or standing group is a dangerous situation. This scenario is defined by a sudden increase in the speed histogram, simultaneous with a change in the histogram of pairwise motion patterns from walking or standing towards diverging and running interaction. Furthermore, the corresponding connected component splits up into smaller connected components or into single pedestrians. The scenario *Brawl* is an uncoordinated movement of a group in which several persons could possibly be fighting. Therefore, features indicating uncoordinated movement are used to define a brawl, such as a high standard deviation of the speed histogram and many different entries in the histogram of the scalar of the motion vectors.

4. EXPERIMENTAL RESULTS

We show experimental results for our complex event detection approach on the basis of a new UAV dataset. In the next section, the UAV dataset is described in detail. Afterward, the experimental results are presented.

4.1. Dataset

The evaluation of our complex event detection approach is based on a new UAV dataset which was captured from an AscTec Falcon 8 octocopter. An UAV provides high mobility and is an ideal platform for surveillance. The images were taken from a height of 85m with a Panasonic DMC-LX3 camera, resulting in a ground resolution of about 1.5cm. The frame rate of the dataset is 1Hz, the images are robustly aligned using Harris feature points. The UAV dataset contains more than 10 different scenarios representing group behavior in different complexity levels in a controlled setting. The pedestrian group in the dataset consists of up to 18 persons who were instructed in advance. The information given to the persons was reduced to a minimum in order to preserve natural behavior. Common scenarios include parallel walking or running in a group. More complex scenarios include the interaction of a moving or standing group with passing individuals or the interaction between two groups. Further scenarios additionally incorporate interaction with obstacles or objects, such as a bottleneck (Figure 5) or a corridor. Additionally, dangerous scenarios with brawling or escaping groups are available. Table 1 gives a detailed summary of scenarios.

For the evaluation of the complex event detector we use reference trajectories which were generated by manually tracking the colored hats in each frame of the image sequences. However, our approach is also able to deal with possibly incomplete automatically generated tracklets because the graph can straightforwardly deal with changing topology. Furthermore, pedestrians within a group usually behave like their surroundings and their missing detection does not distort the general group-related features. For the density calculation an intermittent detection of a present person in a group can be buffered such that the density can be kept over the sequence. This is due to the fact that a single person cannot disappear within one frame. Hence, it is supposed to be still present.

Parallel group motion (walking & running)
Diverging (walking & running)
Converging (walking & running)
Random walking
Individual crossing group (walking & running)
Groups crossing sideways/head-on (walking & running)
Group overtaking group
Group passing wide gap (walking & running)
Group passing small gap (walking & running)
Group passing corridor (walking & running)
Groups passing corridor head-on (walking & running)
Group avoiding obstacle (walking & running)
Groups brawling
Group escaping "bomb" situation

Table 1. List of available scenarios in the UAV dataset.

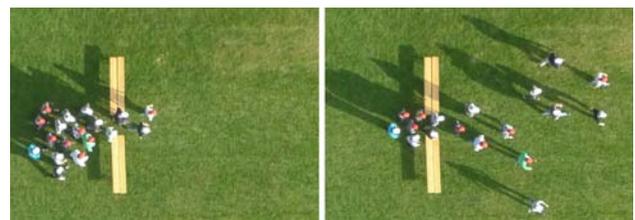


Figure 5. Example taken from our new UAV dataset. Two frames within 4 seconds from the scenario "bottleneck through small gap".

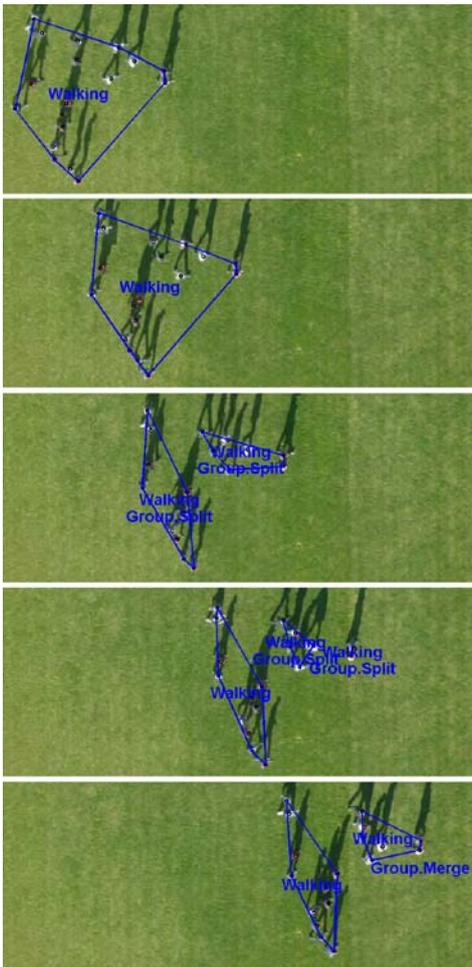


Figure 6. Complex event detection result for the scene “Parallel group motion - walking”. Frames 1, 4, 6, 9 and 12 are depicted (top to bottom).

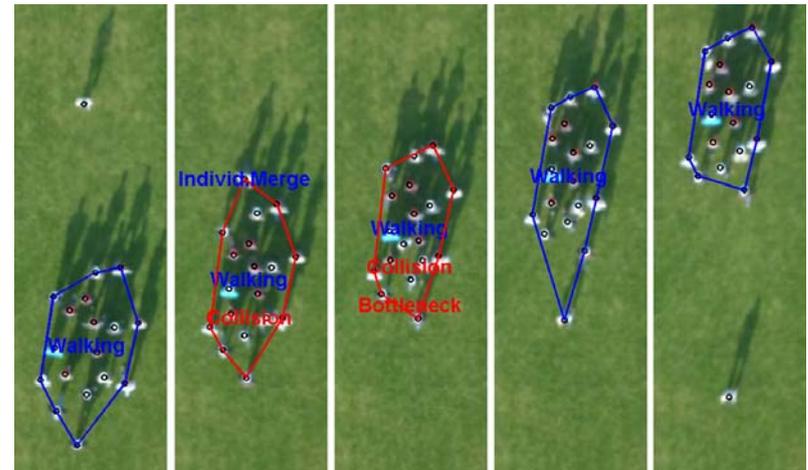
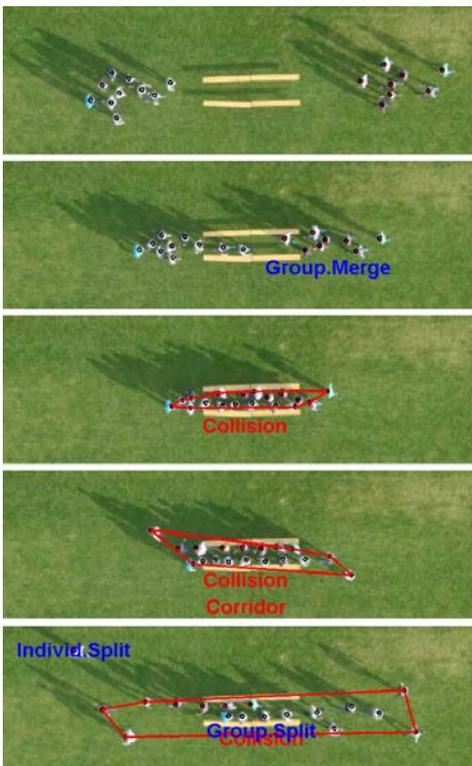


Figure 7. Complex event detection result for the scene “Individual crossing group”. Frames 1, 3, 5, 7, 8 are shown (left to right).

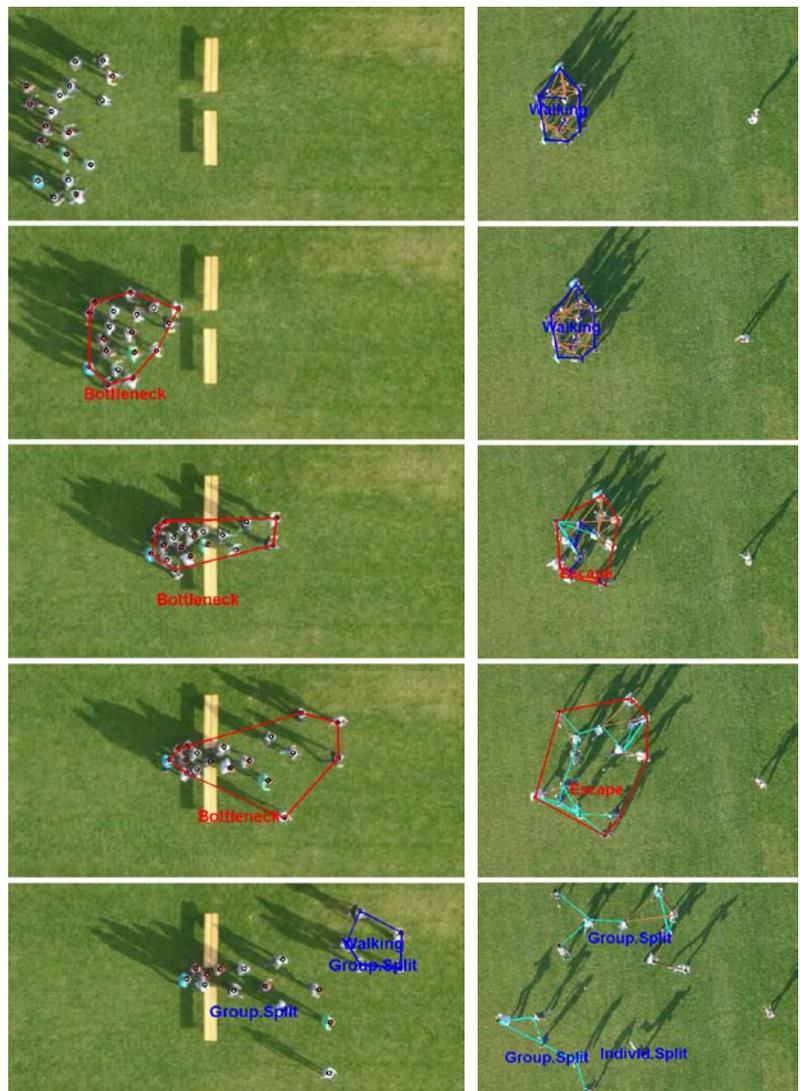


Figure 8 (left). Complex event detection result for the scene “Groups passing corridor”. Frames 1, 5, 9, 13 and 16 are shown (top to bottom).

Figure 9 (top left). Complex event detection result for the scene “Group passing small gap”. Frames 1, 3, 6, 8 and 10 are shown (top to bottom).

Figure 10 (top right). Complex event detection result for the scene “group escaping bomb”. Frames 1, 2, 3, 4 and 5 are shown (top to bottom).

4.2. Complex event detection results

The experimental results of our complex event detection approach are exemplified by some representative sequences from the UAV dataset. The pedestrian interaction graph is depicted only in Figure 10 in order to catch the connected components. The results of our complex event detection algorithm are plotted in blue or red, depending on the potential danger (blue = common, red = danger). Group events are illustrated with a convex hull about the corresponding group.

Figure 6 shows a very common scene, where a group of pedestrians is walking straight ahead and some subsets of the group depart from each other at the end of the sequence. Figure 7 visualizes a walking group which is interfused by an individual head-on from the top to the down. Our algorithm successfully detects the merge of the individual with the group as well as the resulting collision scenario. Figure 8 shows the result for the dataset scenario where two groups are passing a narrow corridor head-on. At first, the group merge is detected (Frame 5), followed by a potentially dangerous collision (Frame 9). Moreover, the formation of lanes is detected (scenario *Corridor*) and finally the groups split again. In Figure 9, a group is passing a narrow gap, resulting in a bottleneck scenario which is successfully detected. Finally, in the last detection result seen in Figure 10, a commonly walking pedestrian group suddenly runs in various directions because a “throwing bomb” scenario was simulated. The scenario *Escape* is successfully detected because of the sudden rise in speed and the decrease of the pedestrian density.

The presented results show that our approach is able to characterize complex group motion on the basis of a predefined model library. Some problems occur because of the construction of the pedestrian interaction graph. In some cases the graph splits up into independent subgraphs and then merges again, even if the corresponding pedestrian group is moving homogeneously, as seen in Figure 6. The reason is that the graph construction is very sensitive to the density calculation. Also, in rare cases false scenario detections are displayed. This may be due to the fact that some scenarios have similar characteristics, although they are uniquely defined.

5. CONCLUSIONS

In this paper, we present a new hierarchical approach for complex event detection in pedestrian groups. The first layer constructs a refined pedestrian interaction graph which analyses low-level pairwise motion patterns between neighboring pedestrians. The second layer assigns high-level group scenarios from a model library to connected components in the graph. This is done by calculating a set of group motion parameters that enable unique identification of the different scenarios independently of the group size. Experimental results are based on a new UAV dataset which comprises more than 10 different types of group behavior. The experiments deliver promising results because meaningful and potentially dangerous group behavior is successfully detected. Future work will aim at the analysis of large crowds in which large connected components have to be split to receive several independent sub-scenario detections. Also, the group motion parameters will be investigated in order to predict potential scenarios. Furthermore, complex event detection results will be used to improve pedestrian tracking algorithms because they provide sophisticated information about group behavior.

REFERENCES

- Burkert, F., Butenuth, M., 2011. Event Detection based on a Pedestrian Interaction Graph using Hidden Markov Models. *In: Stilla, Rottensteiner, Mayer, Jutzi, Butenuth (eds.), Photogrammetric Image Analysis, Lecture Notes in Computer Science* 6952, Springer, pp. 271-283.
- 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. *Proceedings International Conference on Computer Vision Workshops, MSVLC*.
- Dulmage, A. L., Mendelsohn, N. S., 1958. Coverings of bipartite graphs. *Canadian Journal of Mathematics* 10, pp. 517-534.
- Hall, E. T., 1966. *The Hidden Dimension*. Doubleday, New York, 201 p.
- Helbing, D., Molnár, P., Farkas, I. J., Bolay, K., 2001. Self-Organizing Pedestrian Movement. *Environment and Planning B: Planning and Design* 26, pp. 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), pp. 1450-1464.
- Mehran, R., Oyama, A., Shah, M., 2009. Abnormal Crowd Behaviour using Social Force Model. *Proceedings Conference on Computer Vision and Pattern Recognition*, pp. 935-942.
- Mullen, B., Copper, C., Cox, P., Fraser, C., Hu, L., Meisler, A., Smith, C., Symons, C., 1991. Boundaries Around Group Interaction: A Meta-Analytic Integration of the Effects of Group Size. *The Journal of Social Psychology* 131(2), pp. 271-283.
- 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), pp. 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), pp. 831-843.
- Porikli, F., Haga, T., 2004. Event Detection by Eigenvector Decomposition using Object and Frame Features. *Proceedings Conference on Computer Vision and Pattern Recognition Workshops*.
- Rabiner, L.R., 1989. A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition. *Proceedings of the IEEE* 77(2), pp. 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), pp. 1902-1910.
- Zhan, B., Monekoso, D. N., Remagnino, P., Velastin S. A., Xu L., 2008. Crowd Analysis: A Survey. *Machine Vision and Applications* 19(5-6), pp. 345-357.