ROMANIAN JOURNAL OF INFORMATION SCIENCE AND TECHNOLOGY
Volume **21**, Number 2, 2018, 193–205

Analysis of Abnormal Crowd Movements based on Features Tracking

Silviu-Ioan BEJINARIU[1] and Ramona LUCA[1]

[1]Institute of Computer Science, Romanian Academy Iasi Branch, Iasi, Romania
E-mail: Institute of Computer Science, Romanian Academy Iasi Branch, Iasi, Romania

Abstract. Abnormal events in public places can be detected in real time by analyzing the video sequences from video posts, web or surveillance cameras. Sudden changes of speed and/or direction for all people moving in the scene may indicate the occurrence of an abnormal situation and in some conditions also the position in which the event occurred. The detection and tracking of the people in the scene require high computing resources which make this solution difficult to be applied for real time analysis especially in case of crowded scenes. In this paper is proposed an analyzer of scenes participants moving patterns. The *Scale Invariant Features Transform* (SIFT) is used for features extraction and the *Lucas-Kanade* (LK) algorithm for optical flow computation. The main advantage is that it requires low computing resources being suitable for real-time processing. The results obtained by applying the proposed method are compared to those obtained by other approaches presented in literature and also to the subjective evaluation of human observers.

Key-words: Features tracking, Lucas-Kanade algorithm, Scale Invariant Features Transform.

1. Introduction

The abnormal events detection is a challenging task of surveillance systems. Detecting dangerous events that threaten the safety allows quickly alerting intervention teams and taking action to mitigate the effects of these events. Used mainly for crowded areas (public squares, airports, stations, malls) these systems can be used also in small indoor scenes (halls, libraries, markets). Abnormal events can be detected using images, video sequences, and ambient sound if available. The most encountered data sources are surveillance cameras, multimedia content posted on social networks (SN), multimedia content from TV shows and other images or sequences occasionally acquired. The events can be detecting by analyzing the behavior of human participants in the watched area or the movement of other objects, such as vehicles on the roads. The methods involved in the analysis depend on the content and quality of the source images and videos.

Multimedia messages posted on SNs contain mostly person's images in which the upper part of the human body or only the head is visible from frontal position. Face recognition methods [1] are used and facial expression [2], emotion [3] and gestures can be analyzed in this case. The voice can be analyzed also because the sound is clear and noise free. When the video comes from surveillance cameras or multimedia sequences and contain less crowded scenes placed near the cameras, the human bodies can be detected and people moving behavior can be analyzed [4–6]. In this case, recognition of objects in scene may be involved, while it may contain also non-human objects and the real time processing of such scenes may be difficult. The sound may be clear only in case of multimedia sequences. If the surveillance cameras are placed in crowd squares, transportation stations, or shops then the video sequences contain far scenes and the human silhouettes are difficult to recognize even using high resolution videos because of the partial overlays. In this case, the moving objects tracking and trajectory analysis is used to detect the abnormal events. The sound has poor quality and can only be analyzed to detect the appearance of high intensity signals.

Crowd behavior analysis is a topic that has been extensively investigated by researchers in computer vision, video surveillance, human motion recognition, tracking and activity detection for developing applications that automatic detect an abnormal behavior for avoidance of crowd disaster [7–10]. According to [11] there are two types of crowds: stationary crowds and dynamic crowds. In computer vision, crowd behavior analysis has the following approaches: crowd segmentation, crowd dynamic analysis, crowd density estimation [11]. The most robust surveillance systems are based on human actions recognition which requires not only human body detection and motion analysis but also entire scene modelling which are parts of the computer vision domain. The abnormal event detection can be simplified by detecting the type of motion performed by the actors in the scene (walking and/or running). A classification method of human locomotion types from video sequences based on motion parameters clustering is proposed in [12]. The proposed method is able to recognize three different types of movement: walking, jogging and running. More complex systems are based on behavior analysis [13]. In the first step of these systems, the human action recognition is performed and then, the main stages are: (a) behavior representation, which requires features extraction and description techniques to be applied and (b) behavior modelling based mainly on classifications methods [13]. Such a system should either rely on previously known characterizations of normal and abnormal behaviors, or be able to learn these characteristics and self-adapt during surveillance. An abnormal activity detector which is able to dynamically adapt itself to the visual context changes is presented in [14]. The process is fully unsupervised and it is robust to noise in the behavior representation. The Likelihood Ratio Test is used to detect the abnormal behavior. Another unsupervised approach for anomalies detection in video sequences is presented in [15]. The proposed method does not require prior knowledge about the normal or abnormal behavior of the scene participants and the analysis is performed in spatial and temporal context at three levels: point anomaly, trajectory anomaly and multiple object simultaneous anomalies. The trajectory analysis is also used in [16]. The trajectories are described by the control points of the cubic B-spline curves which approximate the trajectories. A dictionary which contains the routes of normal behaviors is built and then it is used to classify the test behaviors. The usage of the optical flow orientations histograms and evaluation of similarity between previously stored frames and new observed frames is proposed in [17]. Using the characteristics of the crowd flow, [18] identifies the following crowd behaviors patterns in visual scenes: bottlenecks, fountainheads, lanes, arches, and blocking. In [7], the theory of topological simplification on the dense field is extended to the sparse particle motion field, which is used to describe the dynamics of the crowd. The global topological structure of the crowd motion is described using the analysis of boundary point structure and extraction of critical point from the particle motion field. Various types of abnormal behaviors, including crowd formation/dispersal, crowds splitting/merging, can be detected by monitoring the changes of the topological structure [7]. A Scale Invariant Feature Transform (SIFT) based mean shift algorithm for object tracking is presented in [9]. The SIFT features are used to find the position of regions of interest in different frames. Other used methods for crowd scenes abnormality detection are: changes of energy-level distribution [19]; evaluation of the whole motion intensity of the crowd which is computed by accumulating all optical flow vectors of a frame and then a threshold is used to detect whether the motion intensity changed suddenly and to decide if there is an abnormal activity [8]; real-time descriptors based on texture analysis for crowd scene dynamics modeling [20]; Support Vector Machine in [21]. More detailed reviews of the used computer vision based techniques are presented in [11], [13] and [22]. Also the sensors networks can be used for events detection. Due to imprecision of data acquisition, also in these cases data processing techniques have to be applied for a correct decision [23, 24]. In [25], a taxonomy of human displacement and scenarios for crowded scene is made for autonomous systems. The possibility of identifying dangerous events using the features extracted from the sound waveforms is analyzed in [26]. Concerning the social networks, it should be mentioned that by analyzing the content of posts, the number of posts about an event can offer information about the occurrence of emergencies and create a detailed description of the situation [27, 28].

This paper proposes a framework for automatic detection of abnormal events based on the analysis of movement in crowded environments using SIFT for features extraction and the Lucas-Kanade algorithm for optical flow computation. In normal situations, the movement is characterized by relatively constant velocity and direction between different points of interest in the scene. Sudden changes of velocity and/or direction for all moving elements in the scene indicate the occurrence of an abnormal situation. The proposed approach analyzes the moving patterns of the features in the scene and then signals the events appearance. The main advantage is that it requires low computing resources being suitable for real-time processing.

The rest of this paper is organized as follows. In the second section the Lucas-Kanade tracking method and SIFT are shortly described and then the abnormal movement detection method is detailed. The experiments described in the fourth section revealed that the SIFT based procedure offers better results in detection of abnormal movements and the last section concludes the paper.

2. Abnormal movement detection

The detection of abnormal events may be accomplished through the analysis of video sequences containing humans which move in a supervised area. Sudden changes of speed and direction for subjects in the scene indicate the occurrence of an event which sometimes can be localized. Abnormality detection can be accomplished by:

- *Trajectory analysis* which is based on object tracking and typically requires an uncrowded environment. The objects trajectories can be approximated by tracking either some distinctive features of the objects or the known models of the objects.
- *Motion analysis* which is suitable for crowded scenes and is performed by analyzing patterns of movement rather than attempting to distinguish objects. It requires computing the optical flow - the vector field which describes how the image changes with time. This

approach has the disadvantage that it requires intensive computation.

The proposed procedure is used to detect the abnormal movements in video sequences acquired using a fixed camera by highlighting sudden increases in the average speed of the detected features. As stable features detector, the usage of SIFT is proposed. The results are compared to those obtained using the Lucas-Kanade algorithm for optical flow computation.

2.1. Scale Invariant Feature Transform – SIFT

In order to detect the velocity of scene's components, a set of stable and distinctive features which are visible in consecutive frames have to be detected. SIFT is a features detector used with good results in pattern recognition, localization, 3D mapping, tracking and image registration, which was proposed in [29]. It is a four steps algorithm: detection of possible features, keypoints selection, keypoints orientation assignment and description generation [30]. First, the set of features candidates is selected by finding the extrema of the Difference of Gaussians (DOG) function computed as the difference of two scaled images separated by a multiplicative factor. In the second step, the most stable and accurately localized features are selected and candidates having low contrast or strong edge response in one direction only are removed. Then, the features orientations are computed using the orientation histogram of local gradients of the closest smoothed image. In the last step, the complete descriptors are computed for all the selected keypoints [30]. The complexity of SIFT depends on the image resolution and number of extracted features which is high for Full-HD images. In [31] is proposed a low complexity version of SIFT using a real time hardware implementation. The video sequences used in our research are low resolution which allows detecting the features in real time.

To compute the moving velocity of the detected features, the correspondences between features in consecutive frames must be found. For each frame, the Euclidian distances between each SIFT descriptor to those detected in the next frame are computed. The computed values are sorted and a match is established when the minimum computed distance is less than 30% of the second distance in the sorted list [30].

Finally, by considering the duration of every frame in the video sequence as time unit, the numeric value of inter-frame velocity (in pixels per frame duration) of each feature is computed as the Euclidean distance between its positions in the two frames (divided by 1). It must be noticed that the moving direction is not considered in this experiment, even if it can offer important information about the position in which an event caused the abnormal activity of the scene's participants.

2.2. Lucas-Kanade algorithm

Optical flow is the apparent motion pattern of objects, surfaces, and edges in a visual scene caused by the relative motion between an observer and a scene. To compute the optical flow with good results, the objects in the scene must not significantly change their color intensity between consecutive frames and neighboring pixels must have a similar motion.

The first version of Lucas-Kanade (LK) optical flow algorithm was developed as an image registration technique [32, 33]. The algorithm examines few potential matches between images with the disadvantage that it can detect the small inter-frame position changes. The LK optical flow computation is based on the search for matches in the spatial intensity gradient in two images. Because it searches for matches in a small neighborhood, the algorithm fails in case of large motions. This issue is solved by applying the same method in scaled images.

2.3. Procedure for Abnormal Movement Detection

The proposed procedure (described in Fig. 1) is based on the evaluation of the average velocity of moving objects present in the scene. First of all, the following aspects should be considered:
- due the image compression, the frames in the video sequences are noisy enough (Fig. 2). For this reason, some wrong correspondences between SIFT features are established which leads to exaggerated values of average distance between features in consecutive frames has some exaggerated values as it is depicted in Fig. 3.a. On the other hand, some fixed features can be detected with small position changes from a frame to another.

- not all detected features belong to moving objects. Because of a large number of this types of detected features, the value of average distance between SIFT features can be influenced.

For these reasons, the proposed procedure includes two filtering stages. The first filter removes all the correspondent SIFT features pairs for which the displacement in consecutive frames is lower than a threshold *thr_move*. The second filter removes the exaggerated values by applying a minimum filter in a sliding temporal window.

```
for each frame in the video sequence
   Convert image to 256 gray levels
   Find the SIFT features
   Find the correspondent SIFT features in the previous frame
   Compute the distance between correspondent features
               positions in the current and previous frame
   Remove correspondent features pairs for which the distance
               is lower than a threshold thr_move
   Compute the average distance between correspondent features
               positions in the current and previous frame
   Add the computed average distance in list avg_dist_list
end for
filter elements in avg_dist_list by applying the minimum filter
               in a sliding temporal window of size wnd_size
the elements of avg_dist_list which are greater than a threshold
               thr_abnormal indicate an increased activity
               in the corresponding frame
```

Fig. 1. Proposed procedure for abnormal movement detection.

The numeric value of the average distance (in pixels) between features position in consecutive frames is equal to the numeric value of the relative velocity of the features considering the duration of each frame as time unit. Thus, in the proposed procedures, the frames in which the relative velocity is greater than a threshold are reported as containing abnormal movements of the scene participants.

To a better evaluation of the results obtained using the SIFT features, a second approach was developed. It is similar to the proposed procedure, with the difference that the Lucas-Kanade algorithm is used to compute the relative velocity of the features detected in the frames of the video-sequence. In the next section, the results obtained by applying the two procedures are analyzed.

3. Experiments and Results

3.1. Dataset description

The experiment was conducted using 6 video sequences (2 for each scene presented in Fig. 2) extracted from "Crowd-Activity-All.avi" file available at "Unusual crowd activity dataset" of University of Minnesota [34]. Some samples that contain both indoor and outdoor video sequences are presented in Fig. 2. In Table 1 the six video sequences are described, with the observation that the frames intervals in which the abnormal activities are present are quite approximate, and based on the observer's subjectivism.

Fig. 2. Sample frames from the test video sequences (from [34])

Table 1. Video sequences used in the experiment

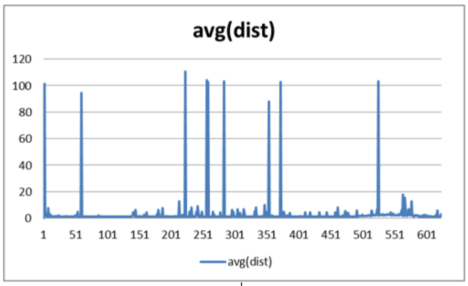
Nr.	Test sequences	Description	#frames	Abnormal activity, frames
1	Seq_1	Outdoor, people walking on grass	625	490-590
2	Seq_2	Outdoor, people walking on grass	825	670-770
3	Seq_3	Indoor, people walking in a corridor	547	320-470
4	Seq_4	Indoor, people walking in a corridor	683	570-670
5	Seq_5	Outdoor, people walking in a square	651	540-640
6	Seq_6	Outdoor, people walking in a square	645	560-640

3.2. Filtering the Measurements

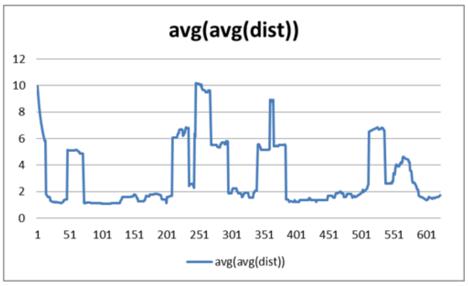
As it was described in previous sections, to avoid measurements values alteration by including the fixed features or detected moving features produced by the noisy frames, an initial filtering step is applied. In step 6 of the proposed procedure the value of the used threshold is *thr_move=1*.

In step 10 of the proposed procedure a minimum filter is applied to remove the exaggerate values produced by wrong correspondences of SIFT features which appear also in noisy frames. The size of the sliding temporal window is *wnd_size=25*. The value of each average distance in frame k is replaced by the minimum value of average distances in frames *[k-wnd_size/2, k+wnd_size/2]*.

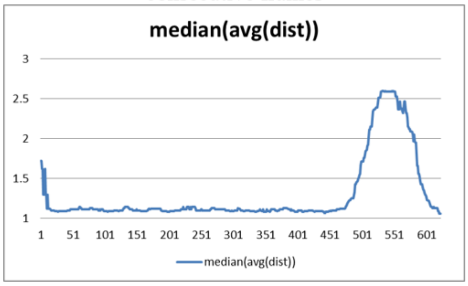
The results obtained by filtering the exaggerate values, for Seq_1 are depicted in Fig. 3. There are eight positions in the sequence where the average distance is about 100 pixels (Fig. 3.a). This is caused by the recording quality and compression method. In fact, for Seq_1, the number of filtered correspondent features in consecutive frames is between 40 and 80, but when the exaggerate values are obtained, the correspondent features number is around 400. Most of these features are placed on the scene's background, with small texture variations which produce wrong associations of SIFT features. Due to the high values, the average filter is not able to smooth enough the evolution of the parameter (Fig. 3.b). Better results are obtained by applying the median filter (Fig. 3.c) but also in this case, in the domain's limits neighborhood, the values of the parameter are not enough smoothed. The best result is obtained by applying the minimum filter (Fig. 3.d).



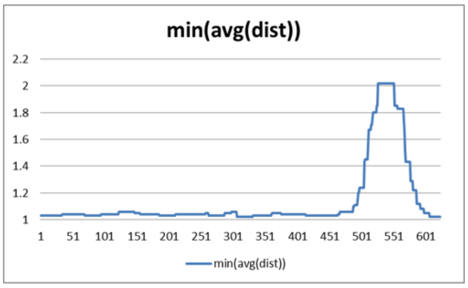
a) Average distance between SIFT features in consecutive frames



b) Result of average filter



c) Result of median filter



d) Result of min filter

Fig. 3. Average distance between SIFT features filtering results, Seq_1 video sequence (horizontal axis: frame number, vertical axis: distances in pixels)

3.3. Results

Finally, in step 11 of the proposed procedure, the filtered velocity array components are compared to the velocity threshold *thr_abnormal* and the frame intervals in which the velocity is greater than the threshold are reported as containing abnormal activities or movements in the scene. In this experiment, the threshold value is *thr_abnormal=1.25* in case of SIFT based motion analysis and *thr_abnormal=1.45* in case of LK motion analysis.

In Table 2 the results obtained using the proposed procedure are summarized.

Table 2. Sequences with abnormal activities in all test sequences

	Sequences with abnormal activities		
Video sequence	**Lucas-Kanade**	**SIFT**	**Observer**
Seq_1	351-355, 491-624	505-580	490-590
Seq_2	279-282, 378-407, 449-453, 479-573, 685-824	699-733	670-770
Seq_3	338-398	346-383, 447-475	320-470
Seq_4	608-643	602-657	570-670
Seq_5	116-126, 558-650	550-624	540-640
Seq_6	319-325, 579-644	582-644	560-640

Considering that the frame intervals which contain an abnormal activity established by direct observation are correct, the results were evaluated using the following statistical measures: *Sensitivity* (true positive rate), *Specificity* (true negative rate) and *Accuracy* (the proportion of true results among the total number of samples) [35]. In this evaluation the abnormal activity detector was considered as a binary classifier of video frames in two categories: abnormal activity and normal activity. The results are presented in Table 3 (all values are rounded to two decimal places).

Table 3. Statistical evaluation of the results

	Lucas-Kanade					
	Sensitivity	**Specificity**	**Accuracy**	**Sensitivity**	**Specificity**	**Accuracy**
Seq_1	0.99	0.93	0.94	0.75	1.00	0.96
Seq_2	0.85	0.74	0.75	0.35	1.00	0.92
Seq_3	0.40	1.00	0.84	0.41	0.99	0.83
Seq_4	0.35	1.00	0.90	0.55	1.00	0.93
Seq_5	0.82	0.96	0.94	0.74	1.00	0.96
Seq_6	0.77	0.98	0.95	0.73	0.99	0.96

By analyzing the content of Table 2 and Table 3, it is obvious that in most cases both procedures detect the abnormal activity intervals. However, some comments should be made:

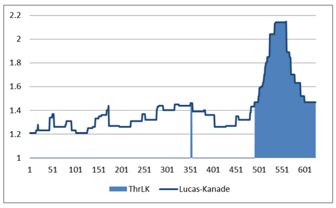
- the Lucas-Kanade algorithm based procedure detects more intervals with abnormal movements in most cases (Fig. 4.a, c, i, k). Most of these have a small duration, less than 10 frames, representing about 1/3 seconds and can be invalidated using this criterion. This is also signaled by the *Specificity* values which are lower than those obtained in case of SIFT version.

- A longer false-positive result is reported by LK based procedure for sequence Seq_2 (Fig. 4.c). It can be explained by the light intensity changes during this sequence. In fact, in Seq_2 many movements are reported for the scene's background.

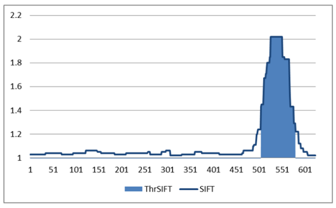
- A false negative result is reported by the SIFT based procedure between frames 384-446 in Seq_3 (Fig. 4.f). This can be explained by the scene's configuration. In video sequences Seq_3 and Seq_4 the viewing angle is rather small while the depth of the scene is quite large. In the mentioned interval the scenes participants are moving away from the camera, making the perceived relative velocity reduced.

- The false negative result reported by LK based procedure for the same sequence is larger. In fact, the movements in the near plane in the last part of the sequence are not detected by the LK based procedure.

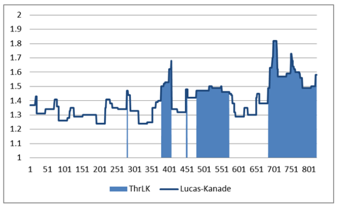
- In most cases, the intervals reported by the proposed procedure are shorter than in case of the human observer's evaluation, but the differences are small enough, about 1/5 seconds. These differences are illustrated by the low values of the Sensitivity measure.



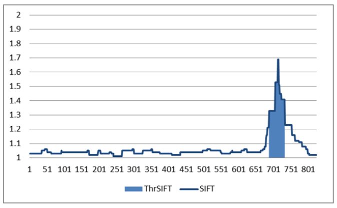
a. Sequence Seq_1, LK detector



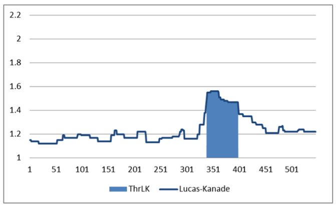
b. Sequence Seq_1, SIFT detector



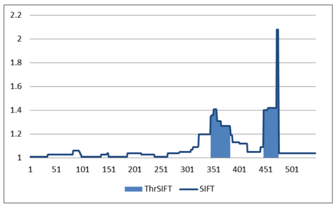
c. Sequence Seq_2, LK detector



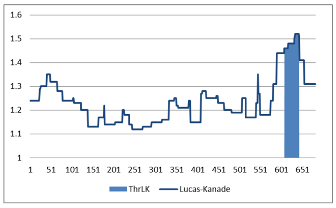
d. Sequence Seq_2, SIFT detector



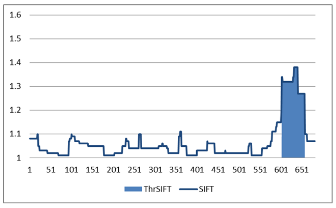
e. Sequence Seq_3, LK detector



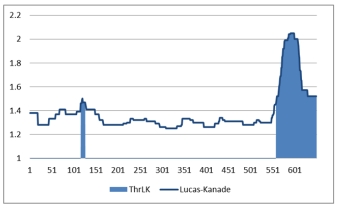
f. Sequence Seq_3, SIFT detector



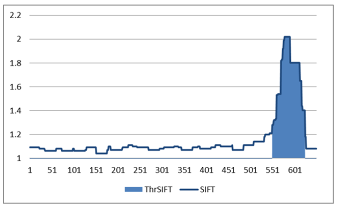
g. Sequence Seq_4, LK detector



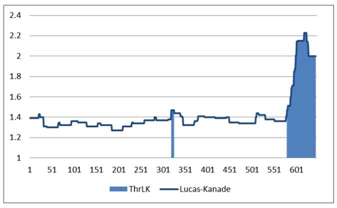
h. Sequence Seq_4, SIFT detector



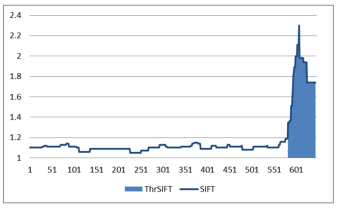
i. Sequence Seq_5, LK detector



j. Sequence Seq_5, SIFT detector



k. Sequence Seq_6, LK detector



l. Sequence Seq_6, SIFT detector

Fig. 4. Sequences with abnormal activities in all test sequences, (horizontal axis: frame number, vertical axis: relative velocity in in pixels per frame duration)

Even the video sequences dataset used in the experiment is reduced, it can be concluded that the SIFT based procedure offers better results than the Lucas-Kanade algorithm based procedure. In fact, with one exception, the values of the Accuracy measure are greater in case of SIFT detector (Fig. 5).

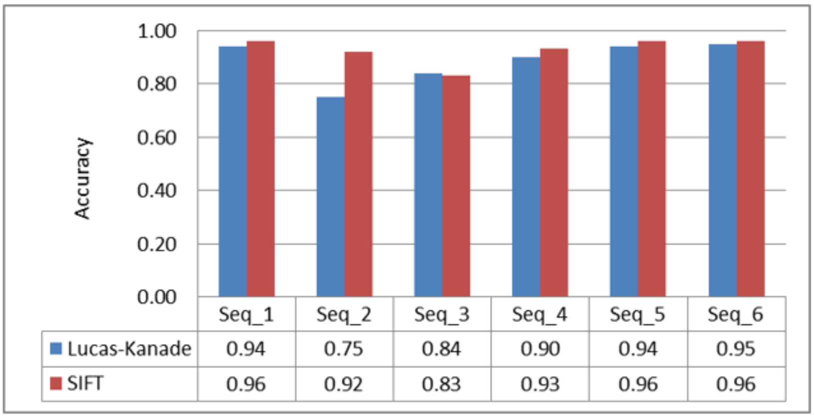


Fig. 5. Sequences with abnormal activities in all test sequences, (horizontal axis: frame number, vertical axis: relative velocity in in pixels per frame duration)

The value of *thr_abnormal=1.25* was chosen by experiments, trying to obtain good results for all tested video sequences. The sequence Seq_2 was also used in [7] and the authors reported an abnormal activity approximately between frames 650-750. This interval is larger than the SIFT based detector obtained, but using *thr_abnormal=1.15*, the results are similar to those reported in [7]. This value is close enough to that which characterizes the normal movement and there is the risk to obtain false positive results in other video sequences. The results obtained for Seq_5 and Seq_6 are similar to those reported in [8], where a dynamic threshold is used to decide the appearance of abnormal movements.

4. Conclusions and Future Work

The article proposes a framework for automatic detection of abnormal events based on the analysis of movement. The main advantage of the proposed method does not require analysis of the scene by detecting, recognizing, classifying and tracking the participants, which are consuming resources. The proposed approaches are based on the determination of the positional changes of SIFT features and optical flow computation using the Lukas-Kanade algorithm, followed by filtering and thresholding which are fast procedures. For the above reasons and also because the analysis is performed on a reduced number of consecutive frames, which does not exceeds one second of delay, the procedures are suitable for real time analysis of video streams.

The video sequences used in the experiments are somewhat similar in terms of visualization angle and distance between camera and scene. So, for other scenes with different illumination conditions or camera position, the procedure's parameters might need to be tuned to obtain good results. This paper describes a work in progress and the authors expect that the procedure can be enhanced. This research will be continued with refinements of the proposed method and human movement analysis in uncrowded scenes which requires automatic detection of body segments in order to model the movement.

Acknowledgements. This work was partially done under the research grant SFPP-984877 "*Modelling and Mitigation of Social Disasters Caused by Catastrophes and Terrorism*", the NATO Science for Peace and Security Programme.

A preliminary version of this research, "*Detection of abnormal movements in crowded scenes*", was presented at the Modeling and mitigation of the effects of the disasters Symposium, 2-4 Nov. 2016, Iaşi, Romania.

References

[1] JOSHILA G.L.K, RESHMI K., *Face recognition in surveillance system*, in: Innovations in Information, Embedded and Communication Systems (ICIIECS), International Conference on, 2015.

[2] CHEN J., CHEN D., GONG Y., YU M., ZHANG K., WANG L., *Facial expression recognition using geometric and appearance features*, in: Proceedings of the 4th International Conference on Internet Multimedia Computing and Service, ACM, pp. 29–33, 2012.

[3] LI J., OUSSALAH M., *Automatic face emotion recognition system, in: Cybernetic Intelligent Systems (CIS)*, 2010 IEEE 9th International Conference on, 2010.

[4] CHEN S., TIAN Y.L., LIU Q., METAXAS D.N., *Recognizing expressions from face and body gesture by temporal normalized motion and appearance features*, in: Image and Vision Computing, 31, pp. 175–185, 2013.

[5] DUFFNER S., ODOBEZ J.-M., Leveraging colour segmentation for upper-body detection, in: Pattern Recognition, **47**(6), pp. 2222–2230, 2014.

[6] HUANG L., TANG S., ZHANG Y., LIAN S., LIN S., *Robust human body segmentation based on part appearance and spatial constraint*, in: Neurocomputing, **118**, pp. 191–202, 2013.

[7] LI N., ZHANG Z., *Abnormal Crowd Behavior Detection using Topological Methods*, in: Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD), 12^{th} ACIS International Conference on, pp. 13-18, 2011.

[8] LIU Y., LI X., JIA L., *Abnormal Crowd Behavior Detection Based on Optical Flow and Dynamic Threshold*, in: Intelligent Control and Automation (WCICA), 11th World Congress on, pp. 2902-2906, 2014.

[9] ZHOU H., YUAN Y., SHI C., *Object tracking using SIFT features and mean shift*, in: Computer Vision and Image Understanding, **113**(3), pp. 345–352, 2009.

[10] WANG B., YE M., LI X., ZHAO F., *Abnormal Crowd Behavior Detection Using Size-Adapted Spatio-Temporal Features*, in: International Journal of Control, Automation, and Systems, **9**(5), pp. 905–912, 2011.

[11] KOK V.J., LIM M.K., CHAN C.S., *Crowd behavior analysis: A review where physics meets biology*, in: Neurocomputing, **177**, 2016, pp. 342–362.

[12] R. LUCA, *Clustering-based Human Locomotion Parameters for Motion Type Classification*, in: Studies in Informatics and Control, **25**(3), pp. 353–362, 2016.

[13] MABROUK A.B., ZAGROUBA E., *Abnormal behavior recognition for intelligent video surveillance systems: A review*, in: Expert Systems With Applications, **91**, pp. 480–491, 2018.

[14] XIANG T., GONG S., *Incremental and adaptive abnormal behaviour detection*, in: Computer Vision and Image Understanding, **111**, pp. 59–73, 2008.

[15] JIANG F, YUAN J., TSAFTARIS S.A., KATSAGGELOS A.K., *Anomalous video event detection using spatiotemporal context*, in: Computer Vision and Image Understanding, **115**, pp. 323–333, 2011.

[16] LI C., HAN Z., YE Q., JIAO J., *Visual abnormal behavior detection based on trajectory sparse Reconstruction analysis*, in: Neurocomputing, **119**, 2013, pp. 94–100.

[17] WANG T., QIAO M. DENG Y., ZHOU Y., WANG H., LYU Q., SNOUSSI H., *Abnormal event detection based on analysis of movement information of video sequence*, in: Optik, **152**, 2018, pp. 50–60.

[18] SOLMAZ B., MOORE B.E., SHAH M., *Identifying Behaviors in Crowd Scenes Using Stability Analysis for Dynamical Systems*, in: IEEE Transactions on Pattern Analysis and Medicine Intelligence, **34**(10), pp. 2064–2070, 2012.

[19] ZHANG X., ZHANG Q., HU S., GUO C., YU H., *Energy Level-Based Abnormal Crowd Behavior Detection*, in: Sensors, **18**(2), pp. 1–16, 2018.

[20] LLOYD K., ROSIN P.L., MARSHALL D., MOORE S.C., *Detecting violent and abnormal crowd activity using temporal analysis of grey level co-occurrence matrix (GLCM)-based texture measures*, in: Machine Vision and Applications, **28**, pp. 361–371, 2017.

[21] DAVY M., DESOBRY F., GRETTON A., DONCARLI C., *An online support vector machine for abnormal events detection*, in: Signal Processing, **86**(8), pp. 2009–2025, 2006.

[22] YOGAMEENA B., NAGANANTHINI C., *Computer vision based crowd disaster avoidance system: A survey*, in: International Journal of Disaster Risk Reduction, **22**, pp. 95–129, 2017.

[23] KAPITANOVA K., SON S.H., KANG K.D., *Using fuzzy logic for robust event detection in wireless sensor networks*, in: Ad Hoc Networks, **10**, pp. 709–722, 2012.

[24] KIM J., PARK S., HAN Y., CHUNG T., *CHEF: cluster head election mechanism using fuzzy logic in wireless sensor networks*, in: Advanced Communication Technology, ICACT 2008, 10th International Conference on, pp. 654–659, 2008.

[25] TEODORESCU H.-N., *Taxonomy and procedures for adaptive autonomous systems in crowded hybrid environments*, in: Proceedings of the 6th International Conference on Electronics, Computers and Artificial Intelligence, ECAI 2014, Pitesti, Romania, 2014.

[26] APOPEI V., *Detection Dangerous Events in Environmental Sounds – A Preliminary Evaluation*, in: Speech Technology and Human-Computer Dialogue (SpeD), International Conference on, Bucharest, Romania, 2015.

[27] TEODORESCU H.-N., *Using Analytics and Social Media for Monitoring and Mitigation of Social Disasters*, in: Procedia Engineering, **107**, pp. 325–334, 2015.

[28] TEODORESCU H.-N., *Emergency-Related, Social Network Time Series: Description and Analysis*, in: Rojas I., Pomares H. (eds) Time Series Analysis and Forecasting. Contributions to Statistics. Springer, Cham, pp. 205–215, 2016.

[29] LOWE, D., *Distinctive Image Features from Scale-Invariant Keypoints*, in: International Journal of Computer Vision, **60**(2), pp. 91–110, 2004.

[30] BEJINARIU S.I., COSTIN H., ROTARU F., LUCA R., NIŢĂ C., LAZĂR C., *Parallel Processing and Bio-inspired Computing for Biomedical Image Registratio*n, Invited Article, in: Computer Science Journal of Moldova, **22**, **2**(65), pp. 253–277, 2014.

[31] SUZUKI T., IKENAGA T., *SIFT-Based Low Complexity Keypoint Extraction and Its Real-Time Hardware Implementation for Full-HD Video*, in: Signal & Information Processing Association Annual Summit and Conference (APSIPA ASC), pp. 1–6, 2012.

[32] LUCAS B.D., KANADE T., *An Iterative Image Registration Technique with an Application to Stereo Vision*, in: Proceedings of Imaging Understanding Workshop, pp. 121–130, 1981.

[33] BAKER S., MATTHEWS I., Lucas-Kanade 20 Years On: A Unifying Framework, in: International Journal of Computer Vision, **56**(3), pp. 221-255, 2004.

[34] Unusual crowd activity dataset, University of Minnesota, http://mha.cs.umn.edu, last accessed on November 3rd, 2017.

[35] FAWCETT T., *An introduction to ROC analysis*, in: Pattern Recognition Letters, **27**, pp. 861–874, 2006 .