Video analytics-based algorithm for monitoring egress from buildings

Maciej Szczodrak · Andrzej Czyzewski

Abstract A concept and a practical implementation of the algorithm for detecting of poten-tially dangerous situations related to crowding in passages is presented. An example of such a situation is a crush which may be caused by an obstructed pedestrian pathway. The surveil-lance video camera signal analysis performed in the online mode is employed in order to detect hold-ups near bottlenecks like doorways or staircases. The details of the implemented algo-rithm which uses the optical flow method combined with fuzzy logic are explained. The experiments were carried out on a set of gathered video recordings from the surveillance camera installed in the campus of Gdansk University of Technology. The results of experi-ments performed on gathered video recordings shows high efficiency of the algorithm.

Keywords Crowd · Crowd behavior · Egress monitoring

1 Introduction

Gathering of a large number of people in a confined area may be the source of dangerous events. Participants of concerts, sport games and other similar ceremonies are imposed to serious physical injuries, and in the worst case they might even lose their life. The emergency situations occurred many times in the history [7, 24, 25]. One of the dangerous situations is a crush which may be caused by the obstructed pedestrian pathway. Obstruction of passages or exits may occur during overcrowded events, where many people are gathered on a small area. Such a situation may take place i.e. in a sport object or an entertainment hall during a football game, a concert, etc. in the moment when every participant intends to leave the building. Regardless of the existence of multiple exits in the building, people tend to choose their known way, for example the way they entered the facility. This may result in rising of a significant slowdown or formation of blockages in places such as passages, halls near door or elevators. Nowadays, monitoring systems are commonly used in such objects. In parallel with

M. Szczodrak (✉) · A. Czyzewski
Multimedia Systems Department, Faculty of Electronics, Telecommunications and Informatics,
Gdansk University of Technology, Narutowicza 11/12, 80-233 Gdansk, Poland
e-mail: szczodry@sound.eti.pg.gda.pl

A. Czyzewski
e-mail: andcz@multimed.org

Springer

popularization of monitoring systems, a continuous process of refining surveillance algorithms can be observed, not only for the video but also often for the audio processing algorithms [12]. Conventional video surveillance systems employ object detection and tracking in order to extract the moving people or car images from the background [4, 5]. Research subjects range from particular algorithms [6, 23] to complete video monitoring systems like systems for event detection in underground stations [13]. The analysis of crowded scenes is much more complex when compared to the non-crowded ones due to problems of detection and tracking of individuals. Various subjects of the video analysis of a crowd are undertaken by scientists. For example: crowd behavior classification results are interpreted as normal or abnormal, by finding the corresponding motion attributes related to opposite movement in a crowd, division, fighting, and others [10, 19, 21]. Another work focus on estimation of a crowd density using diverse methods, like recognition of the head contour using Haar wavelet transform (HWT) and support vector machines (SVM) [27], or image texture statistical analysis [18]. As an alternate example, a system can serve for the detection of overcrowding in underground station platform [16]. In this paper a video analytics algorithm for monitoring egress from buildings or rooms is proposed. The main focus is put on detection of blocking of pedestrian flow near bottlenecks such as door or narrow corridors. The state of pedestrian flow in a given area is determined by examining the rates of movement of pedestrians and their density, obtained by video analysis algorithms. Determination of flow velocity of the crowd does not require detection and tracking of individuals, and even it would be cumbersome to implement. Hence, a method based on optical flow combined with fuzzy logic is utilized for this purpose. The average velocity of flow is calculated in selected points in the frame, according to the direction of the movement of people. An estimation of the state of congestion is performed employing fuzzy logic. For example: a fuzzy rule can be derived like this one: if the observed velocity at checkpoints is decreasing and the occupancy of the area is large, a lock of the area is probable. Automatic recognition of the degree of congestion at critical points allows detecting the lock and responding appropriately, for example, by identifying other exit paths. The processing of the video signal in nearly real time is required to provide a practically usable solution. Therefore, a supercomputer was employed to provide the hardware base for performing the calculations on video streams. The management of the resources of the computing cluster during acquisition and processing of multimedia data is supported by the dedicated platform called KASKADA engineered at Gdansk University of Technology [14].

2 Algorithm

Crowding detection is realized in several stages. In the initial phase *optical flow field* is determined for estimating the speed and the direction of movement of pedestrians. The input for the algorithm performing this task consists of two consecutive image frames. The next step is to analyze the traffic by calculating the average velocity of the stream of pedestrians in each of the previously defined checkpoints. The next step is to determine the occupation of the area by the method of image background subtraction. The obtained information is provided to the input the of fuzzy logic system, whose task is to give the final result of determining the state of congestion of the analyzed area. Illustration of the concept of the detector is shown in Fig. 1. The assumption is made that no other objects are visible than people appearing in the investigated area. Therefore, any additional recognition of object type is not required. The control lines k_i are perpendicular to the exit path and they are defined separately for each camera. The average speed of people v_i obtained by the optical flow method is determined for vectors crossing each control line. The speed vectors v_i are calculated synchronously in each image frame.
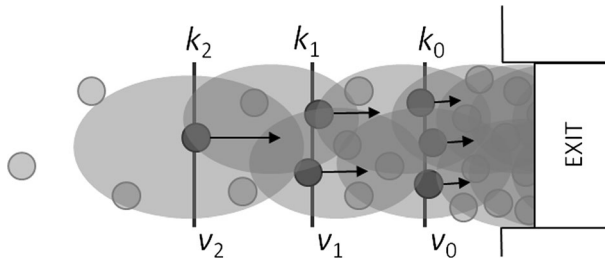
Fig. 1 The concept of clogging detector: k_0, k_1, k_2–control lines, $v_0, ...$–pedestrian flow speed

2.1 Motion detection

The method based on calculation of the optical flow was utilized for detecting of crowd motion speed and direction. The algorithm utilized for obtaining the optical flow field is employing CLG (Combined Local Global) method [3]. Similarly to the characteristic for optical flow algorithms coarse-to-fine strategy, this algorithm uses a multigrid approach, where estimates of the flow are passed both up and down along the hierarchy of approximations. The algorithm combines the advantages of the global Horn-Schunck approach [8] and the local Lucas-Kanade method [17]. Moreover, it was the best-performing algorithm according to the comparison study [1].

The CLG method computes the optical flow field $(u(x,y),v(x,y))^{\mathrm{T}}$ of the image sequence $f(x,y,t)$ at instant t by solving a system of the partial differential equations [3]. The solution is found by the multigrid methods [2, 26]. Typically 4 levels of grid density were utilized, starting from the full image size.

2.2 Pedestrian flow analysis

The obtained continuous flow field (motion direction and velocity determined for each pixel) is sampled in fixed spatial density ($\Delta x, \Delta y$). Vectors extracted in this way intersect with control lines. During the processing of subsequent image frames m and $m+1$, having defined the number of control lines K, we obtain sets of vectors representing instantaneous velocity $v_{m,k,i}$,

where: k—control line number, $k=1...K$,

$i=1... N_k$, N_k—number of vectors which intersect control line k.

Motion velocity at each control line k is calculated according to Eq. (1):

$$v_{m,k} = \frac{1}{N_k} \sum_i v_{m,k,i} \tag{1}$$

The final value of velocity v_k is found as a result of temporal averaging of speed ((1)) in the defined M frames period:

$$v_k = \frac{1}{M} \sum_{m=1}^{M} v_{m,k} \tag{2}$$

The parameter z which represents occupancy of area is obtained with the use of the background subtraction method [9] as defined in Eq. (3):

$$z = \frac{P_{FG}}{P_{TOTAL}} \tag{3}$$

where: P_{FG}—number of pixels not qualified as foreground, P_{TOTAL}—total number of pixels in the image, or more precisely, in the detection area.

Fuzzy logic is employed for making an assessment of the state of pedestrian flow [11, 28]. The parameters determined at previous stages of the processing, namely: velocity and area occupancy constitute the input data to the decision-making system. The Mamadani's method was used as fuzzy inference technique. Membership functions defined for parameters v_k and z defined by Eqs. (2) and (3), named *Speed{k}* and *Occupancy*, have triangular shape. In the discussed case 3, fuzzy sets were used to partition the input space: low (L), medium (M) and high (H). The output fuzzy sets are as follows: none (N), low (L), medium (M) and high (H). The rules, in this example for 3 control lines, are of the form:

IF (*Speed1* **is** {L,M,H}) **AND** (*Speed2* **is** {L,M,H}) **and** (*Speed3* **is** {L,M,H}) **AND** (*Occupancy* **is** {L,M,H}) **then** (*Output* **is** {N,L,M,H}). The membership functions for the discussed system are presented in Fig. 2. For the fuzzy rule inference the fuzzy union and the intersection operators are applied. The centroid method is utilized in the defuzzification procedure.

3 Experiments

The experiments were carried out on the set of gathered video recordings from the surveillance camera installed in the campus of Gdansk University of Technology. Two cameras mounted in the proximity of a lecture hall exit were utilized for gathering the test material. The camera views with indicated crowding detection areas are presented in Fig. 3. The experimental material consisted of 60 recordings. Two types of egress were recorded, namely a normal one, where people flow is fluent and an obstructed one. The efficiency of the algorithm was determined by comparing the algorithm outcomes to the reference data prepared manually by an expert. The recordings content presents people exiting from the lecture hall, whereas the reference data describing the degree of crowding near the door which were calculated for each frame of the video. The degree of crowding can be regarded as a function of two variables as in Eq. (4):

$$R(t) = f(v, z, t) \tag{4}$$

where: v—an average speed of pedestrian flow, z—density of the crowd in the area adjacent to the door.
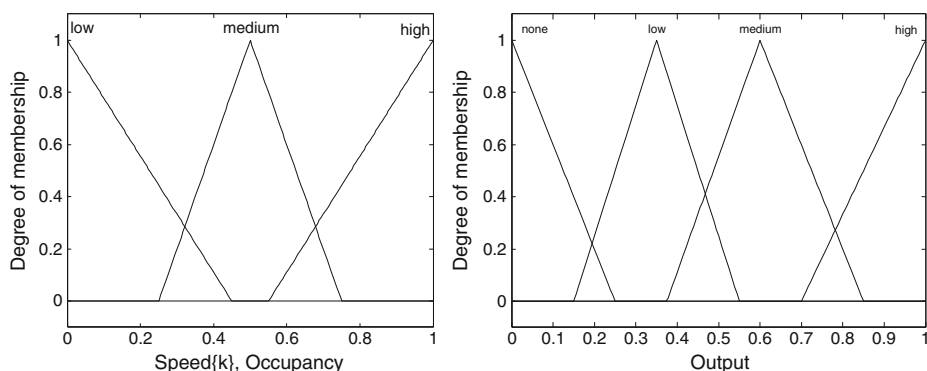

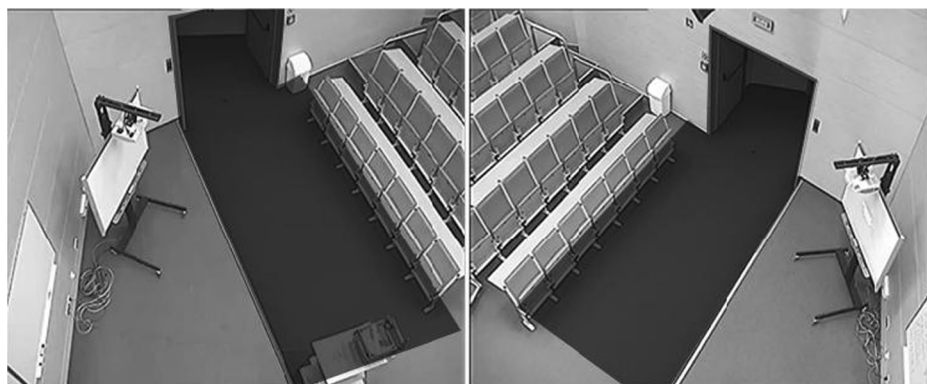**Fig. 2** Membership functions of input and output variables

Fig. 3 Camera view and crowding detection area in Auditorium L (*left*) and Auditorium R (*right*)

The preparation of the reference data by the expert was based on the analysis of the number of people and their movement speed on the way leading towards the exit, according to literature guidelines [20]. The annotation is based on a textual description of degree of crowding categorized as follows: none, low, medium, high. 'None' means normal situation (undisturbed flow), 'high' corresponds to crowded flow. The classification of crowding was made by observation of the number of people and their movement speed in the area adjacent to the exit door. For example, if the number of people was high, which means high density per square meter and their movement speed was low in a specified time period, then the category of 'high' was annotated for this time period. Mathematically each category i can be regarded as a tuple $(L_{i,min}, L_{i,max})$, which describes its lower and upper boundary. For example 'none' is represented by (0, 0.25). The measure of algorithm quality (Q) is defined as the ratio of number of algorithm results (R_t) matching the expert indication (R_{ref}) over time to the total number of results (N):
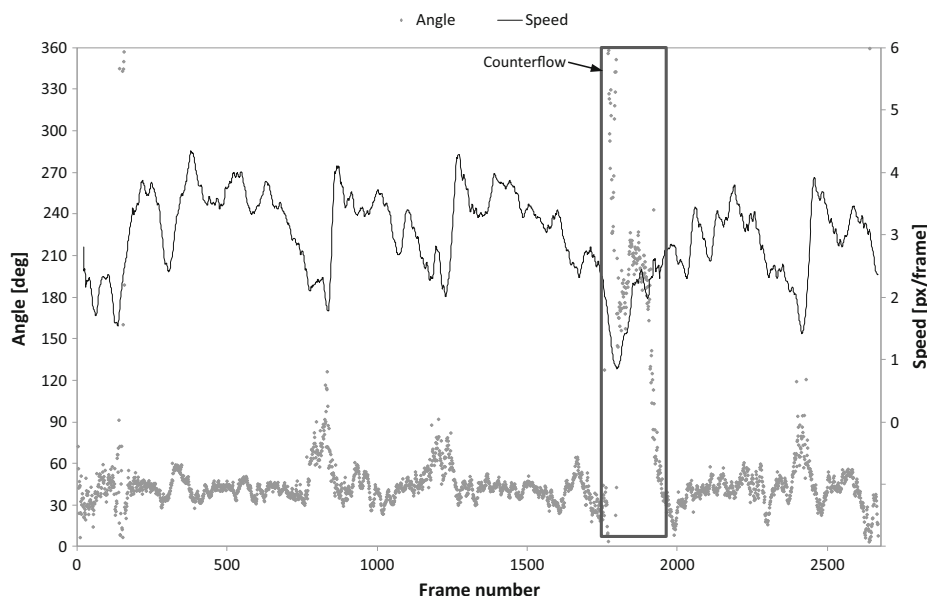

Fig. 4 Average motion speed and mean angle in case of undisturbed flow, Auditorium L
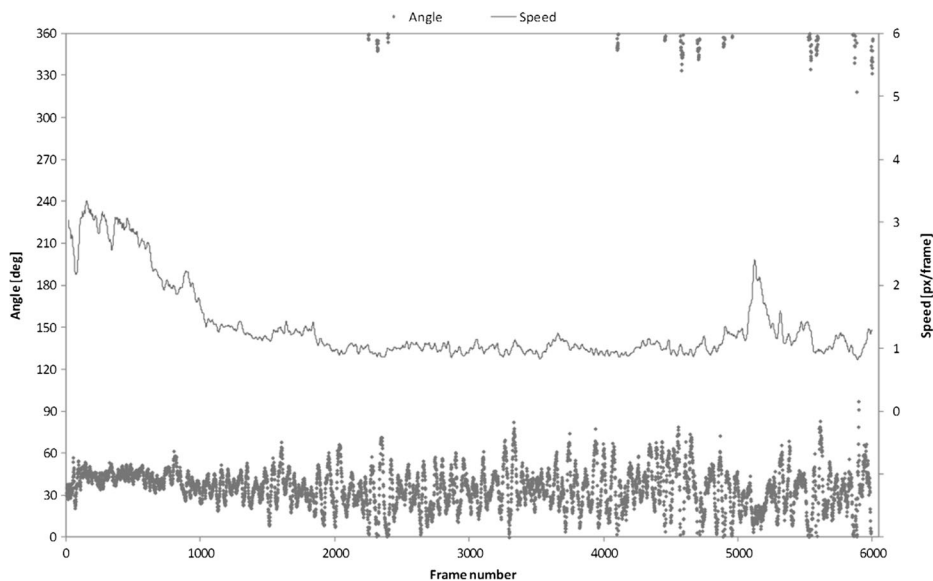
Fig. 5 Average motion speed and mean angle in case of crowded flow, Auditorium L

$$Q = \frac{R_{ref}}{N} \tag{5}$$

where:

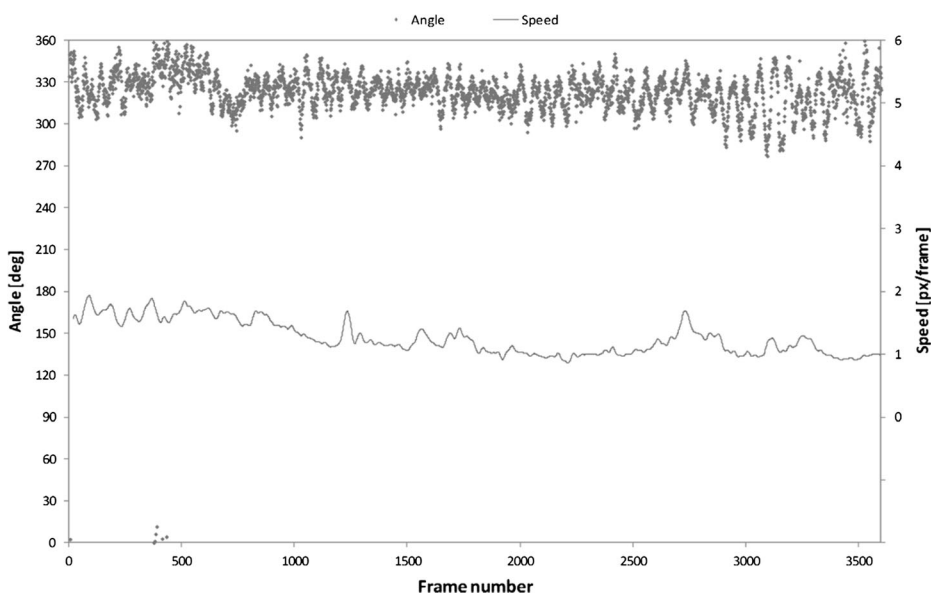$$R_{ref} = \sum_{t=o}^{T} \sum_{i=1}^{C} r_{i,t} \tag{6}$$

Fig. 6 Average motion speed and mean angle in case of crowded flow, Auditorium R
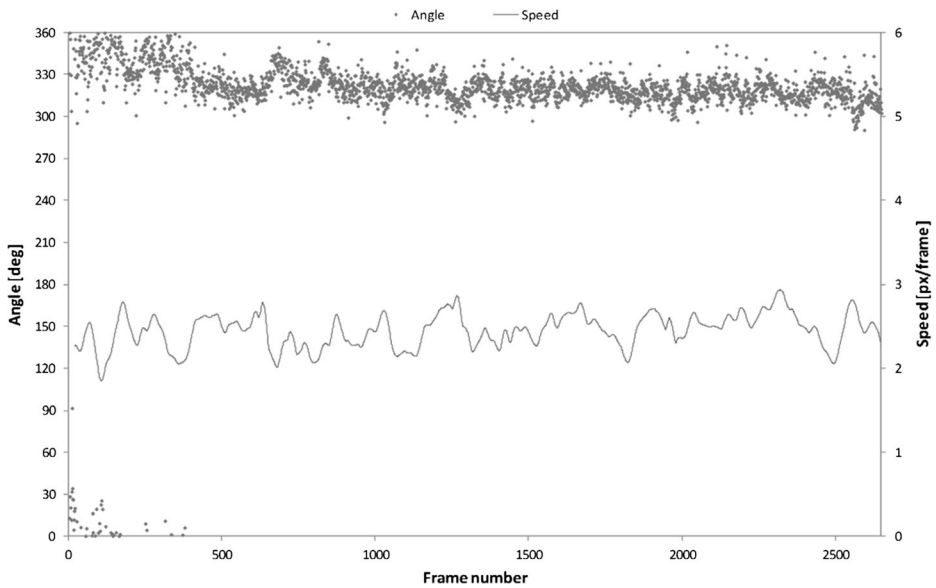
Fig. 7 Average motion speed and mean angle in case of undisturbed flow, Auditorium R

$$r_{i,t} = \begin{cases} 1, & L_{i,min} \leq R_t < L_{i,max} \\ 0, & otherwise \end{cases} \tag{7}$$

C—number of categories, in this case equal to 4.

For a single experiment, N is equal to the number of video frames.

An analysis of pedestrian movement speed and direction was made before testing the algorithm, for the selected recordings which represent undisturbed flow and crowded flow. Movement speed was calculated as a mean of optical flow magnitude in the detection area, and the threshold of 0.5 was applied in order to eliminate insignificant vectors. Movement direction was calculated as a mean of optical flow phase in the detection area:

$$\overline{\phi} = \arctan\left(\frac{1}{n}\sum_{j=1}^{n}\sin(\phi_j), \ \frac{1}{n}\sum_{j=1}^{n}\cos(\phi_j)\right) \tag{8}$$

The result of pedestrian movement speed and direction obtained for undisturbed flow is presented in Fig. 4 and crowded flow case in Fig. 5. We can observe that in case of undisturbed flow motion the speed is about 3 times higher than in the crowded case. The movement direction is uniform in the undisturbed flow case and in the second case fluctuations of the angle can be noted. It is related to commonly observed people swaying at low movement speed and is referred to lateral oscillations [15, 22]. Moreover, in the first 40 s (1,000 frames)

Table 1 Comparison of time-averaged value and standard deviation of angle representing movement direction $\overline{\varphi}$

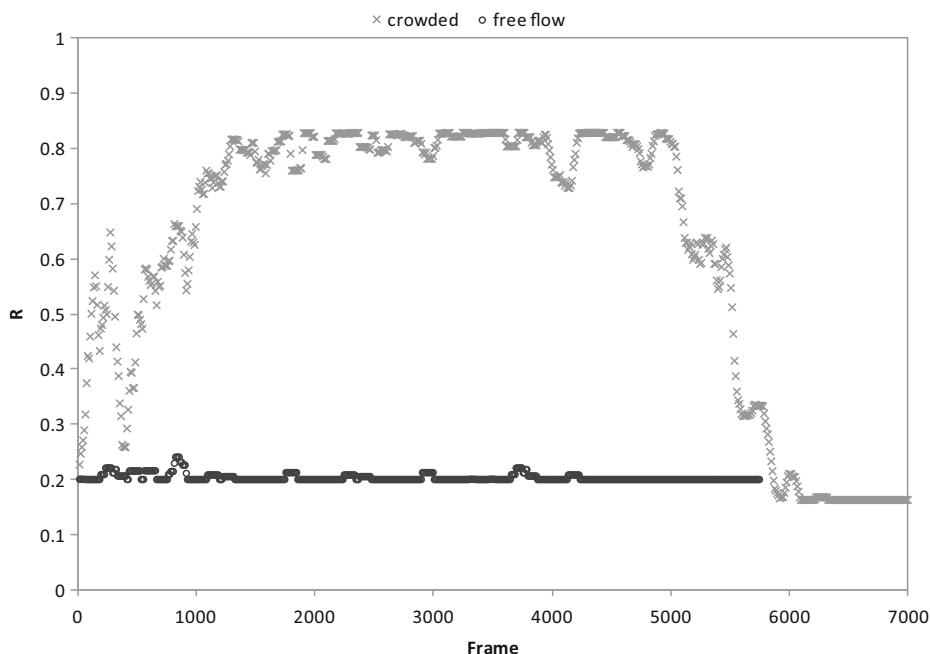|	Fluent, Aud. L[*]	Crowded, Aud. L	Fluent, Aud. R	Crowded, Aud. R
Mean angle [deg]	40.81	34.61	317.24	319.04
Standard deviation	4.36	15.36	8.32	21.47

Fig. 8 Comparison of algorithm output (*R*) for Auditorium L for undisturbed flow (*black*) and for overcrowding (*gray*)

of the crowded case duration (Fig. 5) the pedestrians move fluently, thus a high value of speed and a stable angle value are observed.
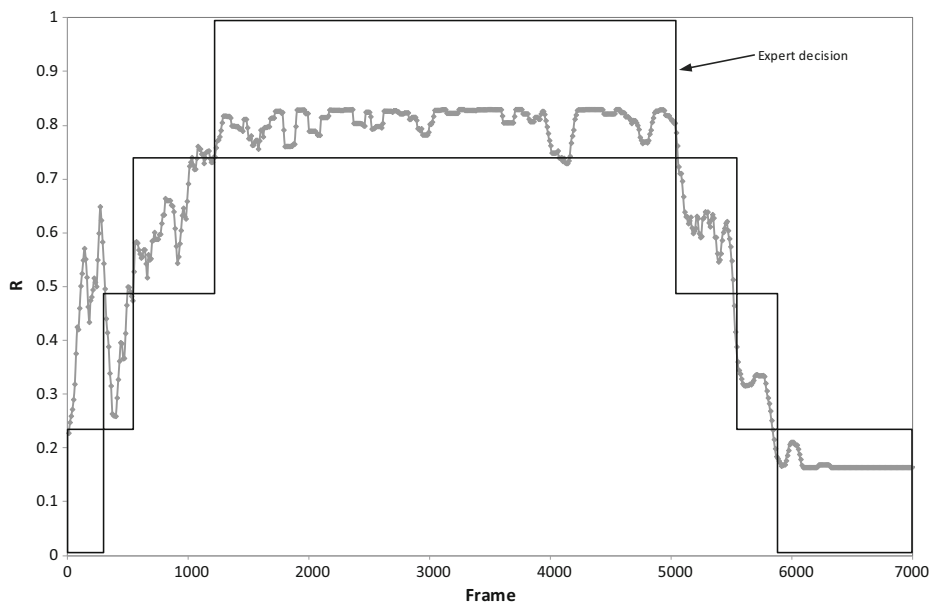

Fig. 9 Comparison of algorithm output (*gray*) of expert decision (*black*) for overcrowding case in Auditorium L

The chart depicted in Fig. 4 represents a continuous fragment of recorded video, therefore we can observe a distortion caused by one person exercising counterflow. Low pedestrian movement speed value in frames between 0 and 180, 747 and 965, 1,170 and 1,280, 2,380 and 2,460 shows time instants in which people were not present in the detection area.

Figures 6 and 7 illustrate obtained movement speed and angle values for crowded and for free flow, respectively, in the Auditorium R. Similar observations can be made as above, nevertheless in case of free flow there is a more coherent value of motion speed than in the sample presented in Fig. 4. This is caused by the continuous pedestrian flow while exiting.

Table 1 presents time-averaged values of the movement direction $\overline{\varphi}$ and the standard deviation for each discussed case. For the undisturbed flow in Auditorium L, a time series where counterflow and movement discontinuities occurred, were not included in the calculation of time average.

The level of crowding (R) obtained as a result of the algorithm application, is shown in Fig. 8. The temporal averaging period was 10 frames (compare to Eq. 2). High crowding is represented by a high value of R and, similarly, low crowding matches a low value of R.

The algorithm quality (Q) was assessed for the discussed cases. An illustration for crowded flow in Auditorium L is presented in Fig. 9. In case of crowded flow Q was 0.94 for Auditorium L and 0.96 for Auditorium R. In case of free flow Q was 0.99 for Auditorium L and 0.95 for Auditorium R.

4 Conclusions

The concept, the implementation and the practical utilization of the algorithm for the detecting of potentially dangerous situations in the crowd were presented. Based on experimental results shown in the paper and on a bigger set of all obtained results, a conclusion can be made that the proposed algorithm is sufficiently effective for detecting the pedestrian crowding near passage bottlenecks. In the future, some enhancements of the algorithm can be done, for example inclusion of the feature of creating statistics of crowding near a particular building exit. Moreover, a connection of multiple cameras to the system is planned in order to enable the pedestrians route prediction function correlated to the current situation in large buildings.

References

1. Baker S, Scharstein D, Lewis J, Roth S, Black M, Szeliski R (2007) A database and evaluation methodology for optical flow. In: Proc. IEEE 11th International Conference on Computer Vision ICCV 2007
2. Briggs W, Henson V, McCormick S (2000) A Multigrid Tutorial, 2nd edn. SIAM Books, Philadelphia
3. Bruhn A, Weickert J, Schnorr C (2002) Combining the advantages of local and global optic flow methods. Vol. 2449 of Lecture Notes in Computer Science, Springer Berlin / Heidelberg, pp 454–462
4. Dalka P (2012) Multi-camera vehicle tracking using local image features and neural networks. In: Proc. 5th International Conference Multimedia Communications, Services and Security, pp 58–67. doi: 10.1007/978-3-642-30721-8_6
5. Dalka P, Szwoch G, Szczuko P, Czyzewski A (2010) Video content analysis in the urban area telemonitoring system. Multimedia services in intelligent environments, smart innovation, systems and technologies 2010, pp 241–261. doi: 10.1007/978-3-642-13396-1_11

6. Hammami M, Jarraya S, Ben-Abdallah H (2011) On line background modeling for moving object segmentation in dynamic scenes. Multimed Tools Appl. doi:10.1007/s11042-011-0935-6
7. Helbing D, Johansson A, Al-Abideen HZ (2007) Dynamics of crowd disasters: an empirical study. Phys Rev E 75:046109
8. Horn B, Schunck B (1981) Determining optical-flow. Artif Intell 17(1–3):185–203
9. Kaewtrakulpong P, Bowden R (2001) An improved adaptive background mixture model for realtime tracking with shadow detection. In: Proc. 2nd European Workshop on Advanced Video Based Surveillance Systems, AVBS01, VIDEO BASED SURVEILLANCE SYSTEMS: Computer Vision and Distributed Processing, Kluwer Academic Publishers
10. Kopaczewski K, Szczodrak M, Czyzewski A, Krawczyk H (2013) A method for counting people attending large public events. Multimed Tools Appl. doi:10.1007/s11042-013-1628-0
11. Kosko B (1997) Fuzzy engineering. Prentice-Hall, Inc., Upper Saddle River
12. Kotus J, Lopatka K, Czyzewski A (2012) Detection and localization of selected acoustic events in acoustic field for smart surveillance applications. Multimed Tools Appl. doi:10.1007/s11042-012-1183-0
13. Krausz B, Herpers R (2010) Metrosurv: detecting events in subway stations. Multimed Tools Appl 50:123–147
14. Krawczyk H, Proficz J (2010) Kaskada—multimedia processing platform architecture. SIGMAP 2010:26–31
15. Liu X, Song W, Zhang J (2009) Extraction and quantitative analysis of microscopic evacuation characteristics based on digital image processing. Physica A 388(13):2717–2726
16. Lo B, Velastin S (2001) Automatic congestion detection system for underground platforms. In: Proceedings of 2001 International Symposium on Intelligent Multimedia, Video and Speech Processing, pp 158–161
17. Lucas B, Kanade T (1981) An iterative image registration technique with an application to stereo vision. In: Proceedings of the 7th International Joint Conference on Artificial Intelligence (IJCAI '81) pp 674–679
18. Marana A, Velastin S, Costa L, Lotufo R (1983) Automatic estimation of crowd density using texture. Saf Sci 28(3):165–175
19. Mehran R, Oyama A, Shah M. (2009) Abnormal crowd behavior detection using social force model. In: Proc. IEEE Conference on Computer Vision and Pattern Recognition CVPR 2009, pp 935–942
20. Polus A, Schofer J, Ushpiz A (1983) Pedestrian flow and level of service. J Transp Eng 109(1):46–56
21. Saxena S, Brmond F, Thonnat M, Ma R (2008) Crowd behavior recognition for video surveillance. In: Blanc-Talon J, Bourennane S, Philips W, Popescu D, Scheunders P (eds.), Advanced Concepts for Intelligent Vision Systems, Vol. 5259 of Lecture Notes in Computer Science, Springer Berlin Heidelberg, pp 970–981
22. Seyfried A, Steffen B, Klingsch W, Boltes M (2005) The fundamental diagram of pedestrian movement revisited. J Stat Mech P10002
23. Szwoch G (2012) Resolving conflicts in object tracking in video stream employing key point matching. In: Proc. 5th International Conference Multimedia Communications, Services and Security, pp 340–349. doi: 10.1007/978-3-642-30721-8_33
24. Taylor P (1989) The Hillsborough Stadium disaster, 15 April 1989: inquiry by the Rt Hon Lord Justice Taylor: interim report, Her Majesty's Stationery Office
25. Wagner U, Fälker A, Wenzel V (2013) Fatal incidents by crowd crush during mass events. (Un)preventable phenomenon? Anaesthesist 62(1):39–46 (in German)
26. Wesseling P (1992) An Introduction to multigrid methods. John Wiley & Sons, Chichester
27. Yin J, Velastin S, Davies A (1996) Image processing techniques for crowd density estimation using a reference image. In: Li S, Mital D, Teoh E, Wang H (eds.), Recent developments in computer vision, Vol. 1035 of Lecture notes in computer science, Springer Berlin Heidelberg, pp 489–498
28. Zadeh LA (1994) Fuzzy logic, neural networks, and soft computing. Commun ACM 37(3):77–84

Maciej Szczodrak Received his M.Sc. degree from the Gdansk University of Technology in 2006. His thesis concerned implementation of algorithms for open air sound propagation analysis. He is interested in audio and video processing, environmental acoustics. He is also concerned about programming. So far he has published, as author or co-author, 30 publications including international journal conference papers and journal articles.

Prof. Andrzej Czyzewski Head of the Multimedia Systems Department is author of more than 400 scientific papers in international journals and conference proceedings. He has led more than 30 R&D projects funded by the Polish Government and participated in 5 European projects. He is also author of 8 Polish patents and 4 international patents. He has extensive experience in soft computing algorithms and sound & image processing for applications among others in surveillance.