Estimation of People Movement in Video Based on Optical Flow Block Method and Motion Maps

H. Chen*,†, R. P. Bohush‡,§, Ch. Chen**, Ch. Chen***, and S. V. Ablameyko‡,∥

† Zhejiang Shuren University, Hangzhou, 310015 China
‡ Polotsk State University, Novopolotsk, 211440 Republic of Belarus
§ Belarusian State University, Minsk, 220030 Republic of Belarus
∥ Joint Institute for Informatics Problems of the National Academy of Sciences of Belarus, Minsk, 220012 Republic of Belarus

e-mail: eric.hf.chen@hotmail.com
e-mail: bogushr@mail.ru
e-mail: ccx0725@126.com
e-mail: ablameyko@bsu.by

Abstract—An algorithm for detecting and tracking moving people on video sequences using the block optical flow method and motion maps is proposed. To reduce time expenditures, a pyramidal representation of the frame and template search are used at the stage of building a preliminary map of motion vectors. The integral optical flow allows one to reduce the resulting amplitudes of the background displacement vectors and increase the resulting amplitudes of the displacement vectors of foreground objects. To improve the accuracy for localization of objects, the additive minimax similarity function is used in the analysis of motion vectors. Objects are tracked based on a modified tracing algorithm using the Kalman filter. The developed algorithm allows one not only to detect a moving object but also to show the trajectory of its movement. The results of experiments are presented that allow evaluating the effectiveness of the algorithm.

Keywords: optical flow, block matching, motion maps, people tracking, movement trajectory

INTRODUCTION

Intelligent vision systems are quite promising for solving various applied problems in manufacturing, medicine, robotics, etc. The construction of such systems is a complex and difficult task that involves obtaining a digital video sequence and processing it in order to extract and then analyze the necessary information. The main stages of the operation of intelligent vision systems include automatic detection and tracking of moving objects under the influence of various kinds of interference and disturbances as well as recognition and description of the actions of objects of interest. Various methods are used to detect and track objects in video images [1, 2].

Tracking the people movement is an important and challenging task. When there is a movement of a person or a group of people, it is necessary to track the beginning of this movement, determine the direction, speed, trajectory, and some other parameters. This behavior can be a sign of a given, for example, an entrance to a controlled area, or an emergency, for example, many people running towards a narrow entrance to a tunnel or building. This task is complicated by the influence of the following factors: change in the illumination of a dynamic scene, the presence of camera noise, change in the shape of an object, the simultaneous presence of several objects with similar characteristics, etc.

The traditional way to determine the behavior of a group of people is to separate the objects of interest from the background and track their movements separately. However, when the group is moving, this method cannot be used due to the numerous occlusions that arise. In recent years, many methods have been used to solve this problem, for example, a convolutional neural network [3] or a social power model [4]. In addition, some methods are known that do not require training. For example, a method is proposed in [5] for identifying five different types of crowd behavior, such as blocking movement, directional movement, movement through a narrowing corridor, movement in a circle and movement from a source based on stability analysis using the Jacobian matrix. A more complete overview of the types of abnormal crowd behavior is given in [6]. In [7], the use of an optical flow was proposed and it was shown that the optical flow and the traffic maps built on it allow one to effectively analyze the behavior of crowds of people, cars on the road, and the movement of the cell population.
Optical flow is understood as a discrete approximation of motion in a three-dimensional scene, carried out by projecting the velocities of three-dimensional surfaces onto an image plane or a visual sensor. The two-dimensional velocity field, which is the optical flow, is used to describe motion in a scene. The optical flow estimates the movement of pixels from frame to frame and is visually represented in the form of displacement vectors for each image point [8]. For this, a shift is found so that a point on the original frame corresponds to a point on the second frame, which determines the length and direction of the vectors. As a rule, the brightness level is used as the characteristics of the image point. The obtained information about the optical flow (speed and direction of movement of neighboring points) is used for spatial segmentation of moving objects. The segmentation accuracy is determined by the efficiency of the optical flow calculation, which depends on the uniformity of the image’s illumination, the structure of objects, and the background’s uniformity.

In this work, we expand the use of the approach proposed in [7] and show that it is possible to track not only the movement of an entire crowd but also each person or a small group of people. Moreover, we show that it is possible to determine the trajectory along which each person moved in a certain period of time. We also propose to calculate the integral optical flow in a more efficient way: based on block matching. Thus, this paper proposes an algorithm for assessing the movement of individuals and groups on video using the block method for calculating the optical flow and motion maps in order to qualitatively control people movement and ensure acceptable computational costs.

1. PERSON MOVEMENT ESTIMATION TECHNOLOGY

The technology for assessing people movement based on the block method for calculating the optical flow and maps of motion vectors is applied to video sequences obtained by stationary surveillance cameras in public places, and it consists in the following.

At the first stage, the optical flow is calculated by the block method, i.e., defines the movement of small moving blocks that are related to dynamic objects in the video. The integral optical flow is then calculated when the optical flow calculations are accumulated for several consecutive frames, which leads to a decrease in the amplitudes of the background displacement vectors and to an increase in the amplitudes of the displacement vectors of foreground objects.

Based on the integral optical flow, it is possible to define and build motion maps that allow describing the movements of blocks in each position together, i.e., to give a statistical analysis of the number and direction of movement of blocks in the direction of each position or away from it.

Regional movement indicators are introduced after that to analyze the movement at the level of areas that consist of moving blocks to analyze the movement of a group of people or a crowd. Crowd movement occurs when many people are moving rapidly in one direction, i.e., there is a directional movement of a group of people, which may indicate an emergency. There are three known characteristics of directional crowd movement [9]: a lot of people move from one area to another, they move fast, and they move in one direction.

The direction of movement in the video indicates the position that people are moving to. When determining the area of heavy traffic, one should take into account the normal speed of their movement, which can vary significantly in different places at different times. For example, people usually walk quickly to the metro during the morning rush hour, but they can slowly wander around the square late at night. After setting the direction and speed of people movement, the area of intense movement can be obtained by means of threshold segmentation [9]. The general scheme is shown in Fig. 1.

2. BLOCK METHOD FOR CALCULATING OPTICAL FLOW

Classic optical flow is a two-dimensional vector field displacement that reflects the movement of pixels between two successive frames. The idea of optical flow is based on two assumptions: pixel intensities do not change when moving between two frames and neighboring pixels make the same movements [8].
To determine the optical flow, two approaches are usually used: differential analysis (gradation methods), which allows one to relate temporal and spatial mismatches, and the block method, which involves finding the best match for fixed-size blocks belonging to different frames. The calculation of optical flow is also performed based on frequency and phase methods.

Modern differential algorithms for tracking features in a video stream are based on the classic works of Lucas–Kanade [8], Tomasi–Kanade, and Shi–Tomasi [10, 11]. The disadvantages of the differential method for calculating the optical flow are significant resource consumption and the complexity of the implementation of algorithms as well as the complexity of accurate numerical differentiation due to the presence of camera noise, low frame rate, and other disturbing influences. Frequency and phase methods for calculating optical flow are even more time consuming. It is assumed that all pixels of the frame block undergo the same movement and the same motion vector corresponds to them. However, in this case, too, the problems of the accuracy of calculating the optical flow and performance are urgent.

The difficulty of accurately determining the displacement of each single pixel is associated with a lack of local information and the influence of noise [12]. Taking into account the property of limiting the constancy of brightness, which allows us to assume that the brightness of the image block should remain practically constant on neighboring frames in a certain neighborhood of the pixel; to calculate the optical approach, it is proposed to select and analyze small blocks of pixels.

The block algorithm for calculating the optical flow assumes frame splitting \( F = f_{xy} \) size \( M_1 \times M_2 \), \( x = 0,\ldots, M_1 - 1 \), \( y = 0,\ldots, M_2 - 1 \), into rectangular blocks of the same size \( b_{ij} \) and search for the corresponding block in the previous frame \( b_{ij}^{t-1} \). In this case, it is assumed that all the pixels of the block undergo the same movement and that they correspond to the same motion vector. Thus, the task of detecting motion on a frame is reduced to the task of finding a motion vector \( v_{ij} \) for each block. In this case, the vectors are determined by the following ratio:

\[
  v_{ij} = \arg \max_{v_{ij} \in O(f(t, i, j, v_{ij})))} (O(t, i, j, v_{ij})),
\]

\[
  O = \{(x, y) | x \in [-u_{max}, u_{max}], \ y \in [-v_{max}, v_{max}]\},
\]

Where \( i, j \) are block coordinates, \( O \) is search area for motion vectors, and \( f(t, i, j, v_{ij}) \) is the function of assessing the similarity of the blocks of the current and previous frames, which takes the maximum possible normalized value equal to one with complete identity of blocks.

According to the expression, a comparison is made for each block of the current frame with the blocks of the previous frame in the specified search area \( O \) and the most similar block is determined by the maximum value of the applied similarity function.

Based on the relative position of the most similar blocks in adjacent frames, the motion vector of the blocks is determined. It should be noted that the found vector of movement for the block may not correspond to its real movement, i.e., will be wrong. The presence of erroneous vectors will not allow performing a high-quality motion estimation.

Erroneous vectors can arise due to the size of the block similarity search area, when, due to the high speed of objects’ movement, the corresponding block goes beyond the boundaries of this area. In addition, the presence of noise and external factors may not allow one to correctly identify similar blocks or lead to an incorrect solution, as a result of which there will be a significant difference between the obtained motion vector and the neighboring ones. Another problem is the high computational complexity for real-time processing in the case of a high-resolution video frame and a large search area.

As a result, a map (field) of motion for blocks is formed, in which the directions of displacement of each block or their absence are shown by vectors at a zero vector for a block.

However, if the time interval between two consecutive frames is very small, then it is difficult to separate the movement of foreground objects from the chaotic movement of the background. The use of an integral optical flow [7, 9] instead of the classical one makes it possible to reduce the influence of the background on estimating the object’s movement and to obtain an area of intense movement.

Integral optical flow is the accumulation of vectors of the optical flow during several subsequent frames. As a result of such accumulation, the resulting amplitudes of the background displacement vectors decrease and the resulting amplitudes of the displacement vectors of foreground objects increase [9]. Thus, it becomes possible to reveal the chaotic nature of the background movement and identify the object’s movement.

The integral optical flow for each block of the image is formed as a result of integrating the values of the optical flow calculated by the method of matching blocks over a given fragment of the video sequence:

\[
  IOFB_{itv}^n (b) = \sum_{i=0}^{itv-1} OFB_{itv}(b_{i+1})
\]

Where \( OFB_b \) is optical stream for video sequence \( I \) calculated by the block method; \( IOFB_{itv}^n \) is integral optical flow; \( itv \) is the interval for calculating the integral optical flow; \( b_{i+1} \) is block position \( f_i (b) \) on the \( (t+1) \) frame of the video sequence \( I \). Furthermore,
In work [7], motion maps are introduced, allowing one to formalize the types of motion of a group of objects. An integral optical flow is used for their formation. These maps were used to analyze and describe motion at the pixel and area level.

In this work, we build motion maps for blocks and areas and use them to identify person movements in video sequences. Thus, a motion map should be understood as an image where the element \( OQB \) in each position shows information about the movement of blocks, the trajectory of which belongs to this position, or the area with the center in this position. Motion maps are first created at the block level, and then regional motion maps can be generated based on them using simple calculations, such as averaging [9].

The basic idea behind motion maps is to record, for each position or area, the block’s movement paths that begin, end, or pass through that position or area. There are two ways to look at motion paths for block-level motion maps. The first way is to consider only the start and end positions. The second is to further examine the positions that have passed along the trajectory of movement. When the first path is used, the motion path is called the simple motion path; when the second path is used, the motion path is called the interpolation motion path. Since region-level motion maps are generated from block-level motion maps, they will also depend on which path is taken.

We created four block-level motion maps. For each position, the first two maps count the number of moving blocks, and the last ones show the total influence of movement for two opposite movement trends, respectively. Based on the approach proposed in [7], we will use the following definitions of motion maps at the block level:

1. **map IQB (number of blocks moving inward):** a map with a scalar value at each position indicating the number of blocks moving to the corresponding position;

2. **map OQB (number of blocks moving outward):** a map with a scalar value at each position indicating the number of blocks moving away from the corresponding position;

3. **map ICMB (complex movement of blocks inward):** a map with a vector at each position indicating the complex movement of blocks moving towards the corresponding position;

4. **map OCMB (complex movement of blocks outside):** a map with a vector at each position indicating the complex movement of blocks moving away from the corresponding position.

On the basis of the four proposed motion maps and vector values, the following variants of the movement of blocks can be calculated according to the method from [7] to assess people movement:

1. positions with the highest values on the map IQB correspond to the positions towards which the largest number of blocks is moving;

2. positions with the highest values on the map OQB correspond to the positions in the direction from which the largest number of blocks moves;

3. positions with the smallest values of the vector modulus on the map ICMB determine the positions, the movement of the blocks to which is the most symmetrical;

4. positions with the smallest values of the vector modulus on the map OCMB determine the positions, the movement of the blocks from which is the most symmetrical.

Taking into account the fact that the search for the most similar blocks on the next frame is carried out in a limited local area relative to the coordinates of the current block on the original frame, it is necessary to take into account the permissible speed of people movement when setting the boundaries of the search area.

### 4. Combined Algorithm for Detecting Moving People

For the possibility of detecting moving objects with an unstable background and reducing computational costs, an algorithm has been developed that is based on calculating the optical flow by the block method using a hierarchical representation of frames for processing and a template method for finding blocks and filtering motion vectors, i.e., a combination of different techniques is used. The algorithm requires the following basic steps:

1. Capture two adjacent frames in size \( N_1 \times N_2 \), \( F^{t-1} \) and \( F^t \) video sequence.

2. Receiving \( p \) levels of the pyramidal representation of frames, in which the pyramid is based on its element-by-element description, and each subsequent description is formed from the previous one by thinning it twice:

   \[
   F'_{xy} = \frac{\sum_{i=-n}^{i=n} \sum_{j=-n}^{j=n} F_{xy}}{k^2},
   \]

   where \( x \in N_1 \), \( y \in N_2 \); \( F'_{xy} \) are pixel brightness of the original image, \( F_{xy} \) is pixel brightness of the reduced image, and \( k \) is reduction factor.
Each subsequent description of a frame, i.e., the next level, in the general case, is the average value of the brightness of the four corresponding readings of the previous level. The hierarchical structure of the presentation of frames is used to reduce processing time. The number of levels of decomposition is determined by the minimum size of moving objects and the size of the frame. In this case, the vector field obtained from the previous iteration is selected as the starting point at each iteration, i.e., each successive iteration refines the vectors calculated at the previous iteration. The main advantages of this approach are reduced processing time, improved noise immunity, and the ability to detect large block offsets. However, there is a high probability of incorrect detection of small objects’ motion due to the use of reduced-size frames at the first stage.

(3) Splitting the image into blocks. The block size is determined by the resolution of the input frame, the permissible speed of movement of objects in the video, and their distance from the video camera.

(4) When calculating the optical flow of the image of the upper level of the pyramid using a template search and building a preliminary motion map for blocks, a rhombus is used as a template. At the topmost level, motion vectors are constructed, and the motion vectors calculated at the previous iteration are refined at lower levels (Fig. 2). The minimax additive function [13] for blocks and on adjacent frames:

$$S_{MMADD} = \frac{1}{NN} \sum_{j=1}^{N} \sum_{i=1}^{N} \frac{\text{min}(b_{ij}, b_{ij}'^{+1})}{\text{max}(b_{ij}, b_{ij}'^{+1})}$$

(5) Clarification of the preliminary movement map using a set of vector candidate of vector $v_{ij}$, which is generated: left vector $v_{i,j-1}$, upper vector $v_{i-1,j}$, upper right vector $v_{i-1,j+1}$; the average value of these vectors [14]; the motion vector obtained from the previous frame. The presence of the last two candidate vectors improves the accuracy of the search for vectors. The methods using candidate vectors are based on the statement that if neighboring blocks belong to the same moving object then their motion vectors are similar. Before calculating the motion information for the current block, a set consisting of the already found motion vectors of neighboring blocks is formed. The generated set of vectors is called a set of candidate vectors. The best vector from the set of candidate vectors is selected as the motion vector in each block.

(6) Performing median filtering of motion vectors blocks within one frame to remove false motion vectors, which involves replacing the motion vector with the vector median of the set. To calculate the vector median from the set, including the current vector and eight neighbors, one is determined for which the sum of the distances to all others is minimal [15]. Distance between two vectors $V_{i_{hk}}$ and $V_{i_{j}}$ is defined as:

$$d = \sqrt{(i_h - i_j)^2 + (j_h - j_j)^2}.$$ 

It should be noted that only nonzero vectors are used when calculating $d$, which makes it possible to prevent the replacement of nonzero vectors by zero ones for the case when most of the neighboring ones are zero vectors. The use of a pyramidal representation of a block for processing leads to errors in determining block movements, especially at the top level. This is due to the fact that the information content decreases when the image is reduced due to the loss of details. The resulting motion map containing the block displacement vectors is used as input data for the lower layer.

(7) Localization of objects. The result of calculating the optical flow is a motion map that displays the movement of areas. In order to correlate areas of motion with objects in the frame, it is necessary to perform the procedure localization of moving objects by determining the characteristics of vectors on the frame taking into account the fact that vectors describ-
ing the movements of one object, as a rule, have similar characteristics.

For localization, we will define the properties of vectors required to select objects on the motion map. Vectors related to one object have the following features: they are located in one area of the frame and form a connected group, are codirectional, and have a similar offset. The passage is carried out on the received movement map. When a nonzero vector is found, mark it as referring to the new object. Check eight neighboring vectors around the found vector. To assess the codirectionality of the vectors, the angle between the motion vectors is determined:

$$\alpha = \arccos \left( \frac{d_{x_1}d_{x_2} + d_{y_1}d_{y_2}}{\sqrt{(d_{x_1}^2 + d_{y_1}^2)(d_{x_2}^2 + d_{y_2}^2)}} \right),$$

where $d_{x_1}, d_{y_1}, d_{x_2}, d_{y_2}$ are the projection of vectors 1 and 2 on the coordinate axis.

The value of the similarity of the magnitude of the displacement of vectors is calculated using the additive minimax similarity function. For each adjacent vector, the deviation angle is checked between the vectors $\alpha$, which should not exceed 90 degrees, and the similarity value should be less than the threshold. If the check is successful, then the current vector, and, accordingly, the block belongs to a moving object.

The algorithm ends after checking all nonzero motion vectors, which makes it possible to detect objects in the video. The objects identified in this way are characterized by the following parameters: the average value of the object’s displacement vector and coordinates and dimensions of the object.

5. DETERMINING TRAJECTORIES OF PEOPLE MOVEMENT

In order to determine the trajectory of an object on a video sequence, it is necessary to establish a correspondence between various objects or their parts in a video sequence as well as to determine other dynamic characteristics of an object or to determine the movement of an object from a given sequence of images. In other words, it is necessary to accompany people.

The trajectory construction procedure is carried out after detecting and localizing moving objects. For this purpose, the following characteristics of the object $O = o_j$ size $W \times W$ are calculated:

(1) The object’s center of gravity $(i_0, j_0)$, which is understood as the center of energy of the light image of an object, is as follows:

$$i_0 = \frac{\sum_{i=0}^{W-1} \sum_{j=0}^{W-1} [o_{ij}]}{\sum_{i=0}^{W-1} \sum_{j=0}^{W-1} [o_{ij}]}, \quad j_0 = \frac{\sum_{i=0}^{W-1} \sum_{j=0}^{W-1} [o_{ij}]}{\sum_{i=0}^{W-1} \sum_{j=0}^{W-1} [o_{ij}]}.$$

(2) The object’s area:

$$S_O = \sum_{i=0}^{W-1} \sum_{j=0}^{W-1} o_{ij},$$

where $o_{ij} = \begin{cases} 1, & i, j \in W; \\ 0, & i, j \notin W. \end{cases}$

(3) Linear dimensions horizontally and vertically, respectively:

$$l_o = \sum_{j=0}^{W-1} [o_j], \quad j_o = \sum_{i=0}^{W-1} [o_i].$$

(4) The vector of motion of the object $v_O$.

The calculated features of localized objects are used to detect them in the next frame. Each object detected and localized using the block method for calculating the optical flow in the first frame, or for the first time in subsequent frames, refers to moving objects. Objects localized on the current frame and detected on the next one using the probabilistic approach are defined as being followed, their characteristics are updated, and the trajectory of such objects is constructed using the Kalman filter [16]. An additive minimax function is used as a similarity function at the maintenance stage.

If an object from the previous frame is not detected in the current frame, then it is referred to as lost. However, in order to be able to detect it on subsequent frames in the event of a short-term loss of optical communication with such an object, the algorithm provides for storing its characteristics for a specified time. The trajectory of the object’s movement is formed as a sequence of its positions, determined by the coordinates of the center, on the frames of the video sequence for a given time.

6. RESULTS AND DISCUSSION

The initial data for the experiments were grayscale and color video sequences obtained in different shooting conditions, seasons, and with different quality characteristics from stationary cameras of video-surveillance systems. Based on the proposed approach, 23 video sequences with a frame resolution of $640 \times 480$ and 22 video sequences with a frame resolution of $1980 \times 1020$ with from one to four moving single people on a dynamic scene and groups of people from three to seven were processed.

The formation of a block movement map is characterized by the maximum computational costs in com-
The experiments have shown that the use of spatial median filtering of motion vectors can significantly reduce the number of false motion vectors (Figs. 3d, 4b); the cumulative effect of the integral optical flow is used to separate the background and the object and obtain areas of intense motion, which are usually areas of interest. In general, this makes it possible to increase the efficiency of the localization of moving people.

The intensity of movement of pixels and their number and direction of movement in a certain area are used to describe the movement and assess people’s behavior. However, for different tasks, threshold val-
ues for traffic intensity and number and direction of movement of blocks should be set depending on the type of scene, camera settings, and tasks to be solved. The parameter of the interval between frames for calculating the integral optical flow determines the threshold of the traffic’s intensity. With its increase, an increase in the intensity of the field movement is observed. Another parameter that determines the threshold value is the size of the area. If the number of people in the group is constant, then a lower threshold should be used for the larger area.

The algorithm provides for storing the main characteristics of a moving object in the event of a short-term loss of optical communication with it in order to be able to detect it in subsequent frames. Determining the trajectory of people movement allows one not only to display their path of movement on a dynamic scene but also, if necessary, find the speed of movement, a sharp change in direction of movement, acceleration, etc.

The proposed approach does not require training; it can be used for monitoring and analyzing the situation as an automatic procedure for monitoring people movement using video surveillance systems.

CONCLUSIONS
An approach is proposed for assessing people movement on a video sequence using the block...
method for calculating the optical flow and analyzing the motion maps generated on the basis of calculating the integral optical flow.

A combined algorithm for calculating the optical flow has been developed, which combines various techniques to reduce computational costs and ensure good efficiency. In this case, the pyramidal representation of the frame is applied. The search for motion vectors for blocks of a frame starts from the upper level of the pyramid, followed by refinement of the results at the lower ones using a template search, in which the coordinates of the points used in the search determined from the center of the region are taken as a template. A rhombus is used as a template, and a set of vectors is used to refine the preliminary motion map, including, relative to the current, left, upper and right upper vectors, a vector representing the average of these three, and a vector from the previous frame.

A distinctive feature of the algorithm is the use of an additive minimax similarity function when matching blocks. To construct motion maps, an integral optical flow is used, which assumes the accumulation of vectors of the calculated optical flow by the proposed combined algorithm during several subsequent frames. When accompanying people, geometric features and motion vectors are used, and the trajectory is constructed using the Kalman filter. The developed algorithm allows not only to detect a moving object but also to show the trajectory of its movement. The efficiency of the algorithm has been demonstrated and confirmed by experimental results.

FUNDING

This work is supported by Public Welfare Technology Applied Research Program of Zhejiang Province (LGF19F0200016, LGJ18F020001 and LGJ19F020002) and the National High-end Foreign Experts Program (GDW20183300463).

COMPLIANCE WITH ETHICAL STANDARDS

Conflict of interests. The authors declare that they have no conflicts of interest.

Statement of compliance with standards of research involving humans as subjects. All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards. Informed consent was obtained from all individual participants involved in the study.

REFERENCES

1. S. Ojha and S. R. Sakhare, “Multiple people tracking using non-iterative tensor voting method,” int. J. Sci. Eng. Res. 6 (6), 1119–1123 (2015).
2. S. Kamate and N. Yilmazer, “Application of object detection and tracking techniques for unmanned aerial vehicles,” Procedia Comput. Sci. 6, 1436–1441, (2015).
3. C. Direkoglu, “Abnormal crowd behavior detection using motion information images and convolutional neural networks,” IEEE Access 8, 80408–80416 (2020).
4. R. Mehran, A. Oyama, and M. Shah, “Abnormal crowd behavior detection using social force model,” in Proc. of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2009) (Miami, FL, 2009), pp. 935–942.
5. B. Solmaz, B. E. Moore, and M. Shah, “Identifying behaviors in crowd scenes using stability analysis for dynamical systems,” IEEE Trans. Pattern Anal. Mach. Intell. 34 (10), 2064–2070 (2012).
6. N. N. A. Sjarif, S. M. Shamsuddin, and S. Z. Hashim, “Detection of abnormal behaviors in crowd scene: A review,” Int. J. Adv. Soft Comput. Appl. 3 (3), 1–33 (2012).
7. H. Chen, S. Ye, O. Nedzvedz, S. Ablameyko, and Z. Bai, “Motion maps and their applications for dynamic object monitoring,” Pattern Recognit. Image Anal. 29 (1), 131–143 (2019).
8. B. Lucas and T. Kanade, “An iterative image registration technique with an application to stereo vision,” in Proc. of the 7th Int. Joint Conf. on Artificial Intelligence (IJCAI’81) (San Francisco, CA, 1981), pp. 674–679.
9. H. Chen, Sh. Ye, O. V. Nedzvedz, and S. V. Ablameyko, “Application of integral optical flow for determining crowd movement from video images obtained using video surveillance systems,” J. Appl. Spectrosc. 85 (1), 126–133 (2018).
10. C. Tomasi and T. Kanade, “Detection and tracking of point features,” in Technical Report CMU-CS-91-132 (Carnegie Mellon Univ., 1991).
11. J. Shi and C. Tomasi, “Good features to track,” in Proc. of the IEEE Conference on Computer Vision and Pattern Recognition (Seattle, WA, 1994), pp. 593–600.
12. D. Fortun, P. Bouthemy, and C. Kervrann, “Optical flow modeling and computation: A survey,” Comp. Vision Image Understanding 135, 1–21 (2015).
13. R.P. Bohush, S.V. Ablameyko, E.R. Adamovskiy, and D. Savca, “Image similarity estimation based on ratio and distance calculation between features,” Pattern Recognit. Image Anal. 30 (2), 147–159 (2020).
14. A. Ishfaq, W. Zheng, J. Luo, and M. Liou, “Fast adaptive motion estimation algorithm,” IEEE Trans. Circuits Syst. Video Tech. 16 (3), 420–438 (2006).
15. J. Astola, P. Haavisto, and Y. Neuvo, “Vector median filters,” IEEE Trans. Circuit Syst. 78 (4), 237–260 (1990).
16. A. Cavallaro and E. Maggio, Video Tracking: Theory and Practice (Wiley, 2011).
Sergey Ablameyko. Born in 1956, DipMath in 1978, PhD in 1984, DSc in 1990, Prof. in 1992. Professor of Belarusian State University. His scientific interests are image analysis, pattern recognition, digital geometry, knowledge-based systems, geographical information systems, and medical imaging. He is a senior member of IEEE, Fellow of IAPR, Fellow of the Belarusian Engineering Academy, Academician of the National Academy of Sciences of Belarus, Academician of the European Academy, and others. He was a First Vice-President of the International Association for Pattern Recognition IAPR (2006–2008), President of the Belarusian Association for Image Analysis and Recognition.

Chaoxiang Chen. Received her MS degree in software engineering from Hangzhou Dianzi University in Zhejiang, China. She is currently a professor and director of teaching quality monitoring center in Zhejiang Shuren University. Her research interest is mainly in the areas of graphics image and intelligent mining. She has completed several researches and published several academic papers in above research areas.

Huafeng Chen. Born in 1982. Associate Professor of Zhejiang Shuren University. Graduated from Zhejiang University in 2003. Received his PhD in 2009 in the field of Earth Exploration and Information Technology at the Institute of Space Information & Technique, Zhejiang University. His scientific interests include remote sensing image processing, GIS application, image and video processing, and multiagent system. He has published more than ten academic articles.

Rykhard Bohush. Graduated from Polotsk State University in 1997. Received his PhD in 2002 in the field of Information Processing at the Institute of Engineering Cybernetics, the National Academy of Sciences of Belarus. Head of Computer Systems and Networks Department of Polotsk State University. His scientific interests include image and video processing, object representation and recognition, intelligent systems, and machine learning.

Huafeng Chen. Born in 1982. Associate Professor of Zhejiang Shuren University. Graduated from Zhejiang University in 2003. Received his PhD in 2009 in the field of Earth Exploration and Information Technology at the Institute of Space Information & Technique, Zhejiang University. His scientific interests include remote sensing image processing, GIS application, image and video processing, and multiagent system. He has published more than ten academic articles.

Rykhard Bohush. Graduated from Polotsk State University in 1997. Received his PhD in 2002 in the field of Information Processing at the Institute of Engineering Cybernetics, the National Academy of Sciences of Belarus. Head of Computer Systems and Networks Department of Polotsk State University. His scientific interests include image and video processing, object representation and recognition, intelligent systems, and machine learning.

Sergey Ablameyko. Born in 1956, DipMath in 1978, PhD in 1984, DSc in 1990, Prof. in 1992. Professor of Belarusian State University. His scientific interests are image analysis, pattern recognition, digital geometry, knowledge-based systems, geographical information systems, and medical imaging. He is on the Editorial Board of Pattern Recognition Letters, Pattern Recognition and Image Analysis, and many other international and national journals. He is a senior member of IEEE, Fellow of IAPR, Fellow of the Belarusian Engineering Academy, Academician of the National Academy of Sciences of Belarus, Academician of the European Academy, and others. He was a First Vice-President of the International Association for Pattern Recognition IAPR (2006–2008), President of the Belarusian Association for Image Analysis and Recognition.