Automatic Detection of the Direction and Speed of Moving Objects in the Video

Ondrej Smirg[1], Zdenek Smekal[2], Malay Kishore Dutta[3], Bhavin Kakani[3]
[1] Department of Telecommunications, Faculty of Electrical Engineering and Communication,
Brno University of Technology, Brno, Czech Republic
[2] SIX Research Centre, Faculty of Electrical Engineering and Communication,
Brno University of Technology, Brno, Czech Republic
[3]Amity School of Engineering and Technology, Amity University, Sector-125 NOIDA, Uttar Pradesh, India.
smirg@feec.vutbr.cz, smekal@feec.vutbr.cz, mkdutta@amity.edu, bvkakani@amity.edu

***Abstract*— This article deals with methods for automatic detection of moving objects in non-standard situations in secure areas, such as nuclear power plants, storage of hazardous materials, etc. Speed and direction, respectively, slow motion to a stop is detected for the monitored object. Specifically: the automatic detection of movement in the opposite direction than normal, passage or move too fast, and stopping or blocking passages. The emphasis is put on real-time processing and on design effective and efficient methods. A new approach is a combination and modification of optical flow techniques and Mixture Of Gaussians method (MOG).**

***Keywords*— Optical flow, Mixture of Gaussian, Farneback, Movement detection, Lucas-Kanade**

I. INTRODUCTION

A frequently solved task in protecting the secured areas or buildings is motion detection of unauthorized or undesirable, animate or inanimate objects. Most of these are special areas where it is necessary to follow certain fixed rules which allow limitations probability of uncontrolled movement; these include the nuclear power plant, guarded buildings, various stores etc. The access and escape corridors, entrances and exits, which provides access to the workplace, and which are secured by any standard identification system for authorization of entry into the field, are in these buildings. These identification systems are well developed now and allow to restrict access and allow it only to truly authorized person. However, situations may arise where there is an accident or an unforeseen situation in some section of the building and authorized persons can behave abnormally. In these situations it is good to have another system that would detect these situations and inform auxiliary units that can intervene and prevent serious situations or crises. The content of this paper is the design of effective automated system that detects abnormal behavior inside the access corridors, exit passageways, etc.

The main parameters, which are taken into account in the design methods, were: real-time processing and maximum efficiency and effectiveness of the methods. As input data sets were created video records with different situations, that represent the correct default behavior of people and other objects in a space and then the records of unauthorized and crisis situations and their combinations (fast motion, wrong direction, etc.). Input camera is equipped with a fish-eye lens type for improving the viewing angle, which increases the area of the monitored area, but the picture is considerably deformed, and it complicates the correct detection. The output of the system should be information what type of illegal situation it is, and position in the image where this situation occurred. Algorithms were designed in other to operate in real-time. It was the main reason why the basic concept of detection methods used optical flow that can work in real-time mode. The system is designed to be modular, so that it can be used in existing systems, and is completely implemented in C + +.

II. SPECIFICATION OF POSSIBLE SITUATION

The situations, which are detected by proposed method, are defined in this chapter. These situations usually occur at some critical events, for example in areas inside nuclear power plants. For these reasons, it is important for the method to work in real-time mode with maximum of efficiency and reliability.

A. Detection of the direction of movement

The first situation, which has to be detected, is the detection of the direction of movement of animate and inanimate objects. This situation concerns the type of input spaces passes to the secure area of building, or vice versa areas of emergency exits. This module is designed to determine the direction of motion of animate and inanimate objects and report situations where these objects move in the opposite direction. Live objects are thinking people and inanimate objects, for example: various transport vehicles (forklifts, electric carriage, etc.).

B. Detection of velocity

This situation describes an event that may occur during an accident in specific area, such as nuclear power plants. In these cases, the staff quickly moved into the area where the accident occurred and tries to solve the problem. On the contrary, other staff is evacuated and tries to leave the building. Detecting this situation should allow the fastest possible identification, facilitating timely execution protection system and initiate rescue forces. In high-risk areas, such as nuclear power plants spaces or hazardous storages, every second is important to allow avert a possible disaster. Except when there is a crash, it can be well planned sabotage, or unauthorized access to protected areas, where the attacker will try to escape from the area.

C. *Detection of stopping objects in protected areas*

Detection of stopping objects in protected areas deal with situations where for example in the areas of emergency exits. There may not be some material, immobilized vehicle or persons, which block the emergency exit. Alternatively, the system can be applied to areas where the risk of leakage radiation and where it is allowed to pass through these spaces only, but it is not desirable that people camped here. This situation is detected when the object is stopped in the area and delays for a period longer than 10 seconds.

III. DESCRIPTION OF THE SYSTEM FOR DETECTION OF THE UNUSUAL SITUATIONS

Since the using of the system is not generalized only to the human figure, but also on other objects, the system does not engage identification of objects, but tries to separate the objects so that the algorithm can determine the properties of each object separately. The image from the camera with fish-eye lens type is used as input data. We used the fish-eye lens type, because we can watch the entire corridor area. The price of camera with a fisheye lens is cost competitive with conventional IP camera. However, it is necessary to use a larger number of conventional IP cameras for monitoring the same area. Input database was created in the test passage, where individual situations and combinations were simulated.

Input data is in motion JPEG format with a resolution of 1280x960 and a very low value of Frames Per Second (FPS), which is 3 frames per second. Preview of the input data is shown in Figure 1.

Figure 1. Input data from camera.

A. *Detection of moving objects*

Detection of objects in the image was divided into several parts. Background modeling techniques were used for the first part of system, specifically the method of Mixture Of Gaussians. This technique creates a statistical background model, which is used to separate moving objects that exhibit statistically different brightness properties than statistical background model. Due to the fact that the model is continuously adapted to the changing conditions, it can filter out processes that are continuously changing in the monitored area (such as rotating fan blades) [1].

The principle of this method is that the recent past of each pixel is modeled using *K* Gaussians and its probability can be expressed by:

$$\left|X_t - \mu\right| < 2.5\sigma. \tag{1}$$

If this does not hold for at least one of Gaussians then the actual pixel is considered as a foreground pixel and the less probable distribution is replaced by the distribution with a mean of X_t, with great variance and low weight ω. For each Gaussian in time t is its weight ω modified by the following formula

$$\omega_{k,t} = (1-\alpha)\omega_{k,t-1} + \alpha\left(M_{k,t}\right), \tag{2}$$

where $M_{k,t}$ equals 1 for every Gaussians where an equality with X_t holds, and for every other it is 0.

Only such a Gaussian that scored equality with X_t needs to have its μ and σ values recalculated in accordance with:

$$\mu_t = (1-\rho)\mu_{t-1} + \rho X_t, \tag{3}$$

$$\sigma^2 = (1-\rho)\sigma_{t-1}^2 + \rho\left(X_t - \mu_t\right)^T, \tag{4}$$

where ρ is given by

$$\rho = \alpha\eta\left(X_t \mid \mu_k, \sigma_k\right). \tag{5}$$

As this computation is very demanding it was simplified as follows

$$\rho_{k,t} = \frac{\alpha_t}{\omega_{k,t}}. \tag{6}$$

It is oblivious that not every Gaussians must belong to the background model. The parameters ω and α matter. The most probable Gaussian is the one that has the highest value of ω and the lowest value of σ. It therefore needs to decide, which Gaussians are included into the background model and which are not.

Often a situation may arise when the clothing of the person or object surface color is not very different from the color, which is represented by one of the Gaussian curves from background, and then there is a situation where the part of object is recognized as the background and is not detected. Another error which may occur in the image is shadow that an object casts, and which has a different color value than the background model. The method used for shadow detection and its removal has very good results [5]. Figure 2 shows an example of detection of a moving object and its shadow.

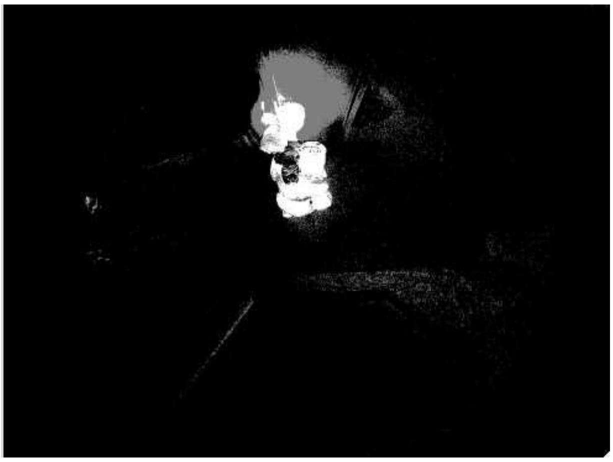

Figure 2. Detection of moving objects by the MoG method with shadow detector(white=object, gray=shadow).

The second algorithm, which in addition to detection of an object can also determine the direction of motion, is an optical flow method [6]. The output of this method is the array direction vectors, the graphical representation can be seen in Figure 3

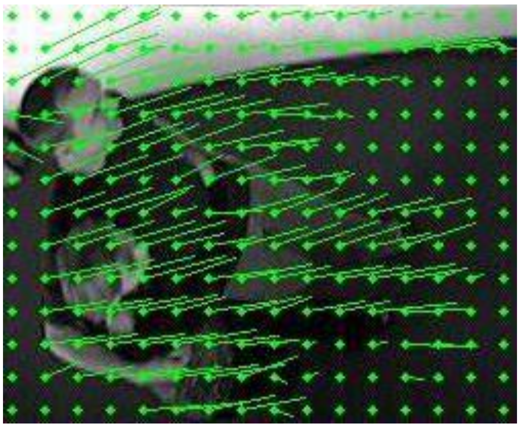

Figure 3. Graphic visualization optical flow method.

Size of the matrix is determined by a predetermined distribution of control points in the image. Each vector gives us the magnitude of change between frames. The matrix contains a large number of zero vectors which are located in areas where there is no movement. Because this method is computationally very intensive and computing power required depends on the size of the matrix, it is advantageous to focus only on non-zero areas. Information about non-zero areas obtained by using the method of MOG, which provides information on areas with moving objects in the form of binary masks. Suitable combinations of both methods suppress their disadvantages. We get rid of the time-consuming methods of optical flow by limiting the area only around moving objects. This combination of methods for MOG removes uncertainty in areas where the color of an object model has very similar properties to the model background color. The figure shows a limited area where the tracked moving object is. The detection is performed by combining optical flow methods in the defined area, which was identified by the MOG.

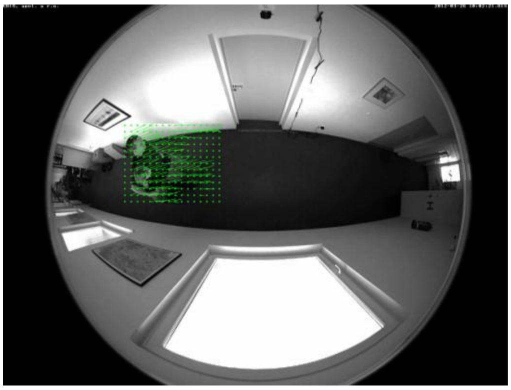

Figure 4. Example of a moving object detection method using optical flow and MOG.

Image distortion by fisheye lens causes deformation of the optical flow at a given point (x, y). We have to recalculate the parameters of the optical flow vectors depending on their position. We know that the corridors walls are approximately parallel. We can use the procedure described in [8] for counting a mapping function. This mapping function is used to correct the direction of movement. We similarly use objects in a vertical direction to derive a mapping function for correcting speed. The mapping function for the direction of movement can be seen in Figure 5. Below we calculate it with the corrected optical flows.

Figure 5. Mapping function curves in the sceme.

Further information, which can be obtained by using the optical flow, is the direction and size of each point matrix. With the help of all of this information direction of the object and its velocity can be calculated. If we come from the assumption that the optical flow of the each point can be described as a set of vectors:

$$k = \left\{ \vec{v}_0, \vec{v}_1, \cdots \vec{v}_n \right\} \quad (7)$$

and the correct direction of motion is given by a unit vector $\vec{w}$. The angle α of the overall direction of the object can be described by the formula:

$$\alpha = \sum_0^i \arccos\left(\frac{\vec{v}_i \cdot \vec{w}}{|\vec{v}_i| \cdot |\vec{w}|} \right) \Big/ i, \quad (8)$$

where i is the number of points that corresponds to the identified object. If the angle is in the range:

$$\left\langle \frac{\pi}{2}; \frac{3\pi}{4} \right\rangle, \quad (9)$$

is evaluated as wrong direction, in other cases it is the right direction. If we express the vector $\vec{v}_i$ using its components in the x and y direction, as v_{ix} and v_{iy}, then the final size k can be determined by the formula:

$$k = \sum_0^i m\sqrt{v_{ix}^2 + v_{iy}^2}, \quad (10)$$

where parameter m takes the values 1 for the vector in the correct direction of movement, and the value -1 for the vector in the wrong direction. Parameter k is the total size of the motion vector of the object. It is not great significance calculate velocity using FPS for practical use. Much more meaningful is the distance covered in a percentage relative to the total monitored area. This can be described so that the distance traveled by the object between the images k is compared with the total distance of corridor in the image. The algorithm is designed so that we can measure the average value of current passage and simulate accelerated movement, which we want detect. With these measurements we can determine the threshold above which can already speed the object considered as substandard.

The last situation which is observed in the video record is the fact that the moving object stays in the area or whether it does not move for an extremely long time. This situation could be solved by using the optical flow. It is possible to set a minimum speed below which speed of an object must not fall. However, a problem arises when the object is stopped completely, because the object is gradual inclusion into the background and we lose information about its position. Our testing also showed another error when there were detecting illegal situation (False Positive), despite the fact that the object was moving normal speed limit. The algorithm has detected a false positive on the outer parts of the image. This situation probably occurred due to nonlinear distortion of image. These situations occurred before the correction of optical flow by mapping function, because the velocity of object was computed lower in the left and right part of images as in the center.

We used modification of the optical flow method for a stationary object detection.[4]. Modified method assumes that the move video content between two frames in close moments is very small and is approximately constant in the vicinity of the reference point P. This method solves cases type uncertainty stopping in the scene for MoG and optical flow method [6], because if the algorithm from [6] detected the decrease of speed below a specified level, then the algorithm [7] finds significant points, which are then tracked. If the total movement of these points does not exceed the minimum speed for 10 seconds, a third type of situation when the object stops in guarded areas is detected. The principle of this detection with found points is shown in Figure 6.

Figure 6. Example of figure caption.

IV. Conclusion

This paper describes a method for the detection of three non-standard situations which may arise in secure areas, and which may be the cause of crisis. Monitored area are detected by special cameras that have fish-eye type of lens, because we want obtain a wide image of corridor or a closed access road. The system was designed to work in real-time mode without using any special hardware support, such as special GPU for increasing computing power of mathematical operations Detection was created to detect both live and inanimate objects in a given space and with maximum efficiency. The output of the method is always mentioned type of illegal situation that occurred and the place or part of the image where this situation occurred. For the detection of moving objects and their direction of motion was used combination of MOG and optical flow methods, which detect individual objects and their movement between frames. Optical flow is calculated for each object separately, according to their detection by MOG. There may be situations where the objects come so close together that they are detected as one. The direction and velocity of an object must be the same (bad or good). If objects are moving in the opposite direction, they will be split and thereby the algorithm detects different directions and speeds. The described algorithm will operate until the HOG is able to separate the objects from each other, regardless of the number of objects. We do not expect to deploy this algorithm in places, where there is so much traffic in normal situation. For the detection of third situation was used optic flow method[7].

ACKNOWLEDGMENT

This research work is funded by projects SIX CZ.1.05/2.1.00/03.0072, EU ECOP EE.2.3.20.0094, CZ.1.07/2.2.00/28.0062, and FR-TI3/170, Czech Republic.

REFERENCES

[1] ZIVKOVIC Z., Improved adaptive Gausian mixture model for background subtraction, *International Conference Pattern Recognition*, UK, August, 2004. pp 28-31**,** Los Alamitos, California,

[2] FIGUEIREDO M. A. T. and JAIN A. K., Unsupervised Learning of Finite Mixture Models, In *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2002, 24: pp. 381–390.

[3] WANG J., Generating daily changes in market variables using a multivariate mixture of normal distributions. In *Proceedings of the 33rd Winter Conference on Simulation*, IEEE Computer Society, 2001, pp. 283–289

[4] LOPEZ DE PRADO, M. (2011), Exact Fit for a Mixture of Two Gaussians: The EF3M Algorithm, The Social Science Research Network. Available at http://ssrn.com/abstract=1931734.

[5] DADGOSTAR F., SARRAFZADEH A., An adaptive real-time skin detector based on Hue thresholding: A comparison on two motion tracking methods, July 2005, *Pattern Recognition Letters 27* (2006), pp.1342–1352.

[6] FARNEBACK G., Two-frame motion estimation based on polynomial expansion, Lecture Notes in Computer Science, 2003, (2749), pp 363-370.

[7] LUCAS, B. D., KANADE T., An iterative image registration technique with an application to stereo vision. In: Proceedings of international joint conference on Artificial intelligence. San Francisco,USA: Morgan Kaufmann Publishers Inc., 1981, s. 674-679.

[8] HUGHES C, DENNY P, GLAVIN M, JONES E, Equidistant fish-eye calibration and rectification by vanishing point extraction.IEEE Trans Pattern Anal Mach Intell. 2010 Dec;32(12):2289-96. doi: 10.1109/TPAMI.2010.159.