VNU Journal of Science, Mathematics - Physics 25 (2009) 143-151

Motion detection and tracking algorithms in video streams

R. Bogush*, S.Maltsev, A. Kastryuk, N. Brovko, D. Hlukhau

Polotsk State University, Blochin str., 29, Novopolotsk, Belarus, 211440

Received 30 June 2009

Abstract. Moving objects detection and tracking in video stream are basic fundamental and critical tasks in many computer vision applications. We have presented in this paper effectiveness increase of algorithms for moving objects detection and tracking. For this, we use additive minimax similarity function. Background reconstruction algorithm is developed. Moving and tracking objects detection algorithms are modified on the basis of additive minimax similarity function. Results of experiments are presented according to time expenses of the moving object detection and tracking.

Keywords: Moving Objects Detection, Tracking, Background Reconstruction , Minimax Similarity Function

1. Introduction

Moving objects detection in video streams is a key fundamental and critical task in many computer vision applications, including video surveillance, as well as people tracking, gesture recognition in human-machine interface, traffic monitoring and so on[1,2]. Detection of moving object should be characterized by some important features: high precision in case of noise components presence on the video streams; flexibility in different scenarios (indoor, outdoor) or different light conditions; efficiency, in order for detection to be provided in real-time.

Basic methods for motion detection in a continuous video stream are: optical flow, frame difference and background subtraction. All of them are based on comparing of the current video frame with one from the previous frames or with background. The most widely adopted approach for moving object detection with fixed camera is based on background subtraction.

For frame comparison of a video information a row of measures are used as unit for measurement of similarity images [4]. Normalized correlation function is widely used among known measures of similarity.

However, the problem of perfection the estimation methods of objects similarity is rather actual, because correlation characteristics of video sequences are far from ideal, i.e., and characterized by a significant level of secondary spikes and main spike inaccuracy [3]. It leads to false identifications of object, or ambiguity of positioning object on the image.

In work [3] attempt to detailed analysis of existing methods for measuring various signal parameters to generate steady against various influences algorithms of objects similarity evaluation is undertaken.

* Corresponding author. E-mail: a.kastruk@mail.ru

The analysis of the considered methods testifies that it is possible to speak only about quasi optimum of the considered similarity evaluation algorithms, depending on external conditions and type of analyzed data. Practically for all methods the basic problem is an accuracy of positioning, which is limited by the base width of the main correlation peak and the presence of intensive level of secondary spikes for the analysis of the image in a mix with noise [3]. Except for this, it is necessary to note computing complexity problems.

In this paper we have introduced effectiveness increase of algorithms for moving objects detection and tracking. For this, we use additive minimax similarity function, which possessing the advanced qualitative characteristic and in comparison with function of normalized correlation, also provides reduction of calculation complexity, as min twice. Background reconstruction algorithm is developed. Moving and tracking objects detection algorithms are modified on the basis of additive minimax similarity function. Also the results of experiments are presented.

2. Minimax similarity function

Functions of similarity are applied for decision of some practical problems in a video processing: moving object detection, object localisation, target tracking, recognition. Normalized correlation function is widely used among known measures of similarity.

In process of algorithms perfection and expansion fields of images processing the correlation coefficient has undergone essential modifications, which have allowed generating on its basis a row of methods measures of similarity differing on properties and characteristics.

In work [4] presented effective family of function similarity for image and video processing. These functions forms an integral similarity estimate based on sequential minimax analysis image elements. In comparison with function of normalized correlation, the minimax function provides reduction of calculation complexity, as min twice. We use an minimax similarity function for decision of some problems: background reconstruction, moving objects detection and target tracking.

Additive minimax similarity function R^s for image $A, N_1 \times N_2$ size, with elements a_{ij} and image $B, N_1 \times N_2$ size, with elements b_{ij} :

$$R^S = \frac{1}{N_1 N_2} \sum_{i=0}^{N_1-1} \sum_{j=0}^{N_2-1} \frac{min\left(a_{ij}, b_{ij}\right)}{max\left(a_{ij}, b_{ij}\right)} . \tag{1}$$

3. Moving objects detection and tracking

3.1. Background reconstruction algorithm

In this section, we have introduced an effective algorithm for background reconstruction. The algorithm takes odd quantity of the frames of input video sequence in which moving objects are present and produced background of the dynamic scene. Frames for processing take out through the set interval. Algorithm includes two basic procedures: calculating binary matrix of motion detection between neighbours work frames and background reconstruction for each of two frames. The constructed images are classified as input data (work frames) for the following iteration of algorithm. Algorithm steps are described as the following:

1. Extraction of N frames of input video streams for vector construction, which includes images of these work frames:

$$w = S^1, S^2, .., S^N = S^k, S^{k+L}, ..., S^{k+(N-1)L}, \tag{2}$$

where $$N \geq 3 \,\&\, N \equiv 1(\mathrm{mod}\,2) \tag{3}$$

L – interval between work frames ($L \in \{20....50\}$ witch guarantee the correct of background reconstruction);

k - number of image frames from N.

2. Testing l for every step:

$$\text{if } l \equiv 0(\mathrm{mod}\,2), \tag{4}$$

where $l \in \{N-1, N-2, ..., 1\}$, then:

2.1 Forming the binary matrix of motion detection using two images S^k *and* S^{k+L} for each RGB color channel separately as:

$$m_{ij}^{q} = \begin{cases} 1, \; if \;\; \dfrac{\min(s_{ij}^{k}, s_{ij}^{k+1})}{\max(s_{ij}^{k}, s_{ij}^{k+1})} \leq T, \\ 0, if \;\; \dfrac{\min(s_{ij}^{k}, s_{ij}^{k+1})}{\max(s_{ij}^{k}, s_{ij}^{k+1})} > T. \end{cases} \tag{5}$$

where T - is a threshold to determine whether the intensity value at the point changes; $q \in \{1, ..., N-2\}$.

The utilization RGB channels improve the accuracy moving object localization.

2.2. Binary image processing of morphological filters. For this purpose we use opening operation:

$$M^{q} = M^{q'} \; X, \tag{6}$$

where X is a structuring element.

3. In opposite case p.2 $l \equiv 1(\mathrm{mod}\,2) \,\&\, l > 1$, producing the vector which includes l intermediate background as:

3.1. Create the vector with elements looks as matrix of motion detection $M^{\{q,q+1\}}$. This matrix includes moving objects for frame of S^{k+1}. Matrix $M^{\{q,q+1\}} = \{m_{ij}^{q,q+1}\}$ can be calculated as:

$$m_{ij}^{q,q+1} = m_{ij}^{q} \cdot m_{ij}^{q+1}. \tag{7}$$

3.2. Forming a vector of the work background. Background is defined as result of removing each pixel of moving objects from frame of S^{k+1} and paste of pixels of background from frame of S^k for this area. We extract the moving object from frame using (8):

$$s_{ij}^{k} = \begin{cases} s_{ij}^{k}, if \; m_{ij}^{k} = 1, \\ s_{ij}^{k+1}, if \; m_{ij}^{k} = 0. \end{cases} \tag{8}$$

4. Steps 2–3 of the algorithm are repeated. The procedure is terminated after (N-1) steps.

5. Background update.

5.1 Deleting the first frame from a vector w and produce cyclic shift for each frame shift to the left on one position.

5.2. Extract new frame from video sequence applying interval L and use this image as a position $S^{k+(N-1)\cdot L}$ of a vector w.

5.3. Steps 2–4 of the algorithm are repeated till l =1.

To simplify the description, we use a group of schematic diagrams (fig.1).

Background reconstruction is in practice just the starting video processing step in a system that is usually supposed to work in real-time. Therefore, it is important to make this step time efficient. In figure 2 time expenses are resulted by background reconstruction for iterative algorithm, background information fusion algorithm and Gaussian mixture background model for 23 sequences. All experiments are implemented on a personal computer (CPU - AMD Athlon (tm) 64 2200 Mhz, RAM - 960Mb) for different scenarios of indoor and outdoor surveillance.

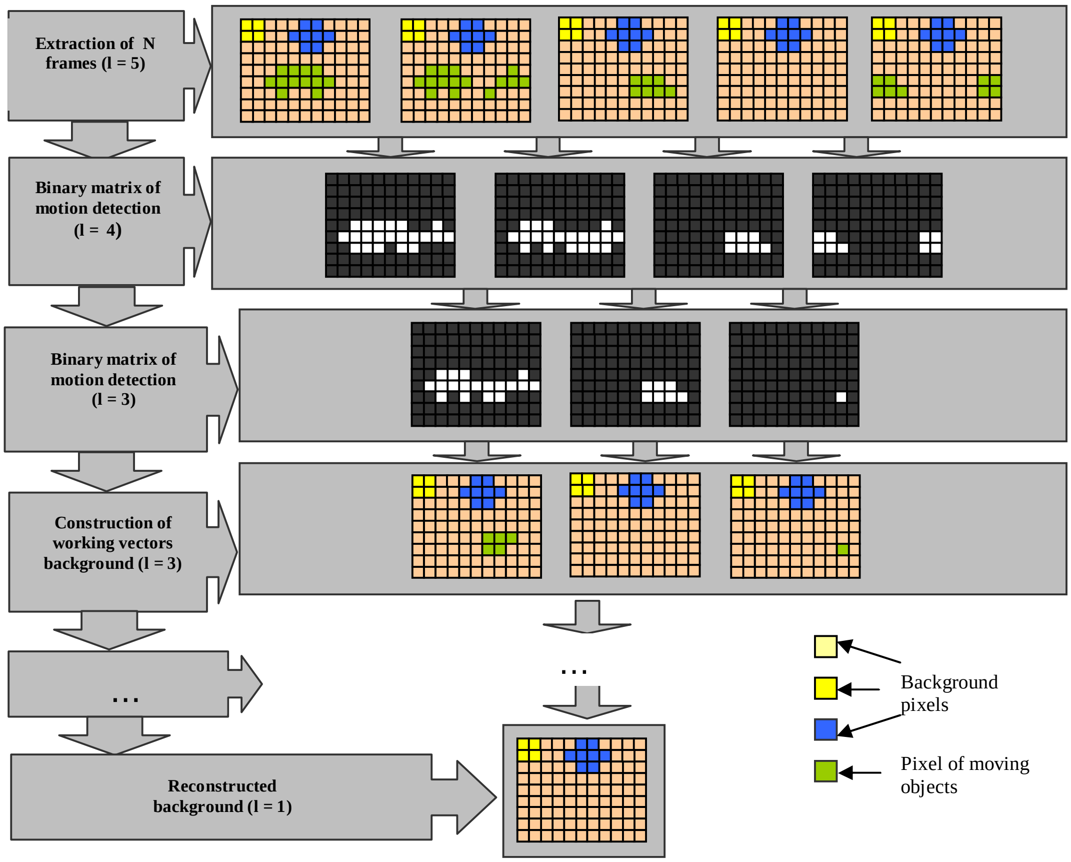

Fig. 1. Schematic diagram for background reconstruction.

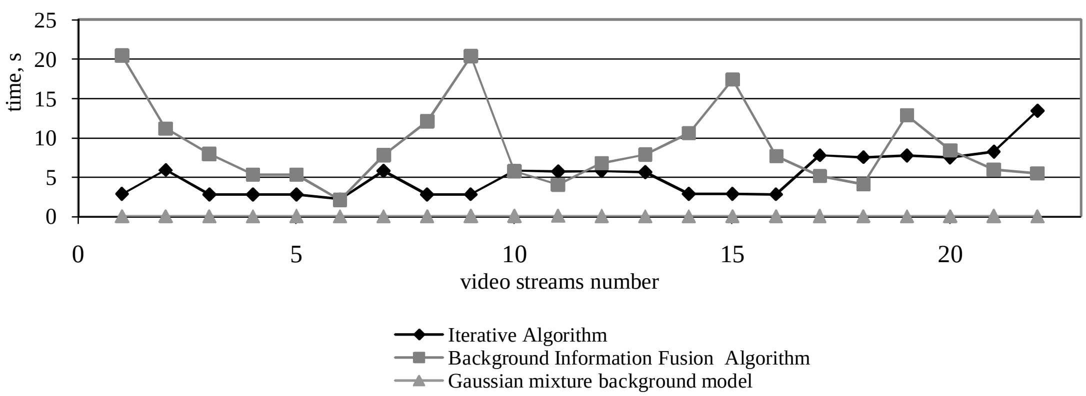

Fig. 2. Time analysis.

Figure 3 shows some examples for background reconstruction from the benchmark suite of our video sequences. On figure 4 results of motion detection for four sequences are presented.

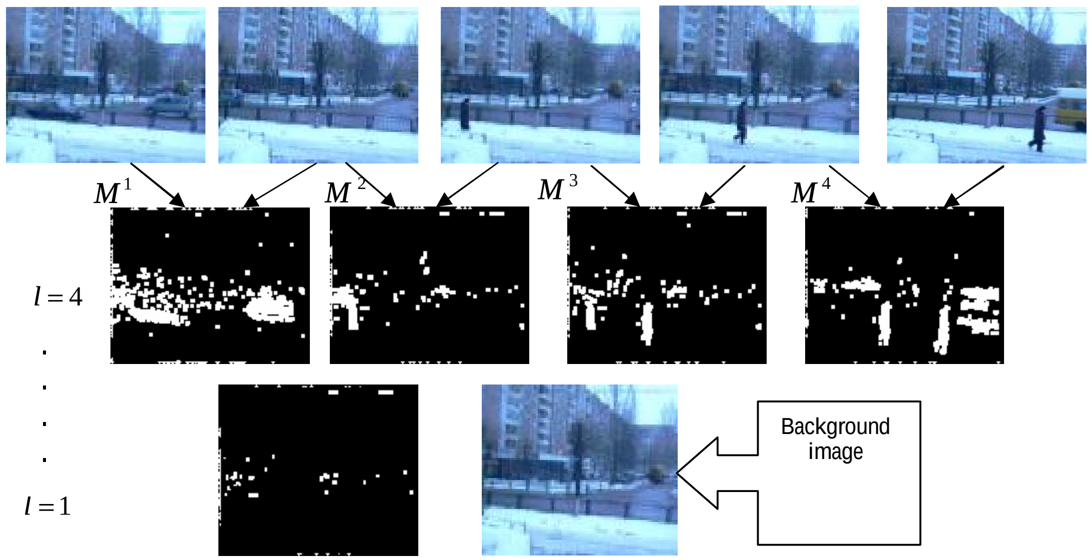

Fig. 3. An example of background reconstruction.

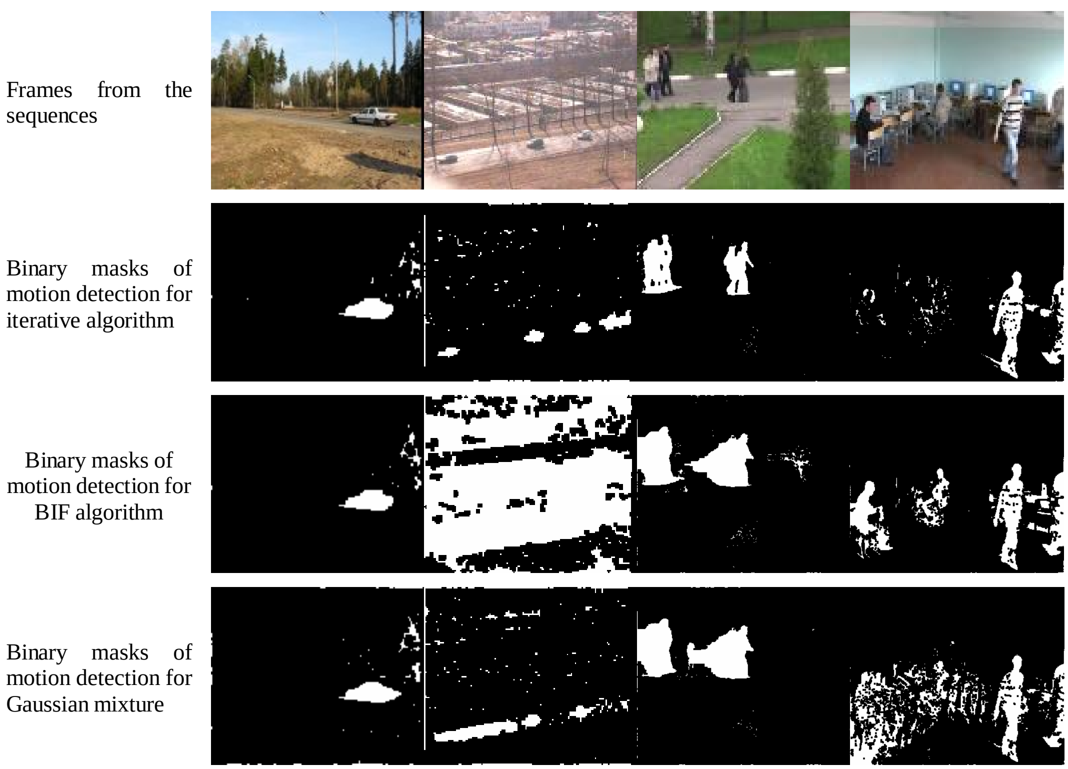

Fig. 4. The motion masks for several sequences.

High quality of the image of background providing thanks an optimum choice of parameters of N, k, T, L.

For satisfaction criteria of «quality /computational complexity» optimal N to be chosen from $3 \le N \le 11$.

For moving objects and case of approximately matching of speed, number of work frames can be chosen minimal. However, it is necessary to estimate speed of moving objects. For the control of moving automobiles the parameter k to should be chosen from $30 \le k \le 50$.

Parameter k can be defined more precisely if speed of moving objects is known. For moving objects with different speed, number of work frames can be chosen maximal

3.2. Moving objects detection

Moving object detection aims at segmenting regions corresponding to moving objects such as vehicles and humans from the rest of an image. Detecting moving regions provides a focus of attention for later processes such as tracking and behavior analysis because only these regions need be considered in the later processes. We use technique based on the background subtraction, that uses level of similarity for comparison of corresponding fragments of the video sequence adjoining frames. If the similarity function does not exceed the preset threshold T_R, the decision on presence of changes for analyzable fragment of the frame, is made (fig.5). For do it, we use an minimax similarity function (1).

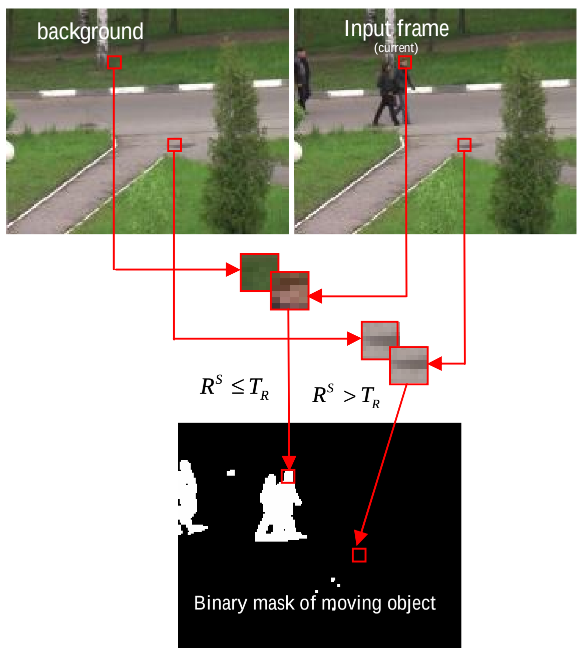

Fig. 5. Schematic diagram for moving object detection.

After applying one of these approaches, morphological operations are applied to reduce the noise of the image difference:

- morphological erosion:

$$S(-)B = \left\{ C \in Z^2 \middle| \forall \ \mathrm{b} \in \mathrm{B,c+b} \in \mathrm{S} \right\}, \tag{9}$$

where S – image, B - structuring element 5×5;

- morphological open:

$$S \circ B = (S(-)B) \oplus B, \tag{10}$$

where B - structuring element 3×3;

- morphological dilatation:

$$S \oplus B = \left\{ C \in Z^2 \middle| \exists s \in S, \ \mathrm{b} \in \mathrm{B:c=s+b} \right\}, \tag{11}$$

where B - structuring element 5×5;

Figure 6 shows some examples for moving objects detection. Small regions in masks of motion detection, are eliminated with a morphological operations, the other foreground pixels are segmented into motion regions by a connected component algorithm. (fig.6c).

a) b) c)

Fig. 6. Moving cars detection: a) original picture; b) binary mask of moving object; c) after morphological processing.

3.3. Moving objects tracking

Moving objects tracking requires to match regions detected in two (or more) consecutive frames. In real video, the matching has to deal with false detections due to noise and to errors with objects in the scene which stop and resume moving, or may become partially occluded. Therefore matching the detected regions in order to derive a trajectory requires an appropriate representation of the detected regions and a similarity function to match these regions.

We use modification of algorithm [6] for effectiveness increase moving objects tracking. For this, we use additive minimax functions similarity (1) allowing with a high degree of accuracy to process a video information for moving objects tracking. In comparison with function of normalized correlation, the offered minimax function also provides reduction of calculation complexity, as min twice.

We apply the following modification algorithm based on additive minimax similarity function. Given several motion windows at frame t, the corresponding motion windows at frame t+1 have to be found. The search of corresponding windows is done in two steps: for each motion window at time t, the window with the greatest of similarity function is searched in frame t+1. Each window with the highest similarity function, matching window, found in frame t+1 has to be validated as a region corresponding to a moving object in the same frame. Given a window, look for its closest translate in frame t+1, assuming that no transformation except translation can occur between two successive images. Examples of the trajectories for two cars are shown in figure 6. In figure 7 time expenses are resulted by one frame (640×480) processing using normalized correlation and additive minimax similarity function for 20 real sequences.

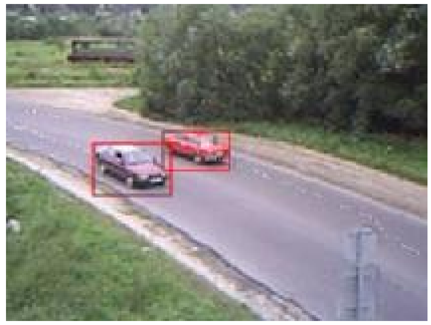

Fig. 7. Tracking of several cars.

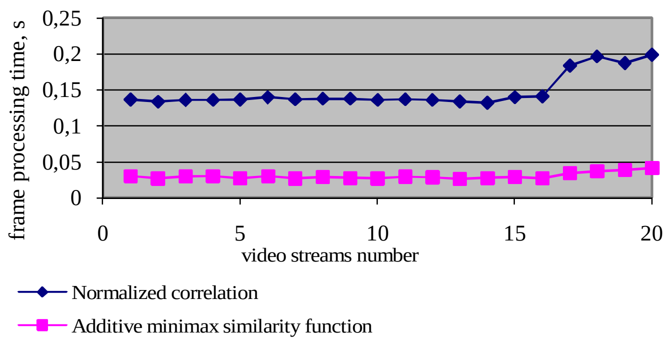

Fig. 8. Time costs.

4. Conclusion

We have presented in this paper effectiveness increase of algorithms for moving objects detection and tracking.

For this, we use additive minimax similarity function, which possessing the advanced qualitative characteristic and in comparison with function of normalized correlation, also provides reduction of calculation complexity, as min twice. Background reconstruction algorithm is developed. Moving and tracking objects detection algorithms are modified on the basis of additive minimax similarity function.

The efficiency of our approach is illustrated and confirmed by our experimental videos.

Reference

[1] J. Ferryman, S. Maybank and A. Worrall. Visual Surveillance for Moving Vehicles, *Int. Journal of Computer Vision* 37 (2) (2000) 187.

[2] P. Kumar, A. Mittal and P. Kumar. Study of Robust and Intelligent Surveillance in Visible and Multimodal Framework, *Informatica* 32 (2008) 63.

[3] S. Chambon, A. Crouzil. Dense matching using correlation: new measures that are robust near occlusions. *Proceedings of British Machine Vision Conference,* Norwich, Great Britain, vol. 1 (2003) 143.

[4] R.Bogush, S. Maltsev. Minimax Criterion of Similarity for Video Information Processing. *Proceedings of IEEE International Conference Control and Communications (SIBCON 2007)*, Tomsk, April 20-21 (2007) 120.

[5] Y. Chen, C. Han, X. Kang and M.Wang, "Background Information Fusion and its Application in Video Target Tracking" *Proc. of 7th Int. Conf. on Information Fusion*, Stockholm, Sweden, June 28 to July 1 (2004) 747.

[6] G. Baldini, P. Campadelli, D.Cozzi and R. Lanzarotti. A simple and robust method for moving target tracking, *Proceedings of the International Conference Signal Processing, Pattern Recognition and Applications (SPPRA2002)*, Crete, Greece June 25 – 28 (2002) 108.