Adaptive Moving Object Segmentation Algorithms in Cluttered Environments

Yuriy Ivanov[1], Dmytro Peleshko[2], Oleksandr Makoveychuk[3], Ivan Izonin[4], Igor Malets[5], Natalia Lotoshunska[6], Danylo Batyuk

1. PIT Department, Lviv Polytechnic National University, UKRAINE, Lviv, S. Bandery street 12, E-mail: yuriy.ivanov@email.ua

2. PIT Department, Lviv Polytechnic National University, UKRAINE, Lviv, S. Bandery street 12, E-mail: dpeleshko@gmail.com

3. PIT Department, Lviv Polytechnic National University, UKRAINE, Lviv, S. Bandery street 12, E-mail: omakoveychuk@gmail.com

4. PIT Department, Lviv Polytechnic National University, UKRAINE, Lviv, S. Bandery street 12, E-mail: ivanizonin@gmail.com

5. PMITT Department, Lviv State University of Life Safety, UKRAINE, Lviv, Kleparivska street 35, E-mail: igor.malets@gmail.com

6. PIT Department, Lviv Polytechnic National University, UKRAINE, Lviv, S. Bandery street 12, E-mail: yuriy.ivanov@email.ua

Abstract – **This article considers two algorithms which are designed for taking into consideration such events as the light changing in the video frame, the background micro-movements and occlusion of an object. These events determine the negative influences on the segmentation process, because they often arise in the eyeshot of surveillance cameras that are operated outdoors.**

Keywords – **segmentation of moving objects, negative impacts, mixture of normal distributions, selective background updating**

I. INTRODUCTION

Segmentation of moving objects and background is one of the key stages of the analysis of digital video streams. At this stage, should take into account such factors as illumination changing, partial occlusion of objects, the presence of moving shadows, noise digital surveillance cameras etc. Therefore, the process of segmentation requires to create and periodically update the background model, from which depends the correctness all subsequent stages of processing video stream, and computing resources which required this processing. Overall, all designed segmentation methods can be divided into the following groups:

- methods of processing background;
- probabilistic methods;
- methods of time difference;
- optical flow methods.

In probabilistic methods model of the background is created by modeling of the pixel process, i.e., for each pixel change intensity between the processable and the previous frame is considered as a linear combination of one-dimensional normally distributed random variables. More perfect are the methods which creating model of pixel by pixel for the entire frame, which uses a mixture of normal distributions separately for the background and moving objects. Based on the time of existence and dispersion of Hausian in each mixture, it is possible to determine which of them belong to the background. The pixels whose values do not belong to the background distributions are considered to belong to a moving object before will appear hausian, which allow a sufficient accuracy to segment them into the background. Such approach allows taking into account the slow changes in the light of video frame by the settings of parameters of Hausian.

The time difference method allows segmenting moving objects and background by calculating the difference between the pixels of two or more consecutive frames. Obviously, these methods are determined dynamically changing of the background, but usually can't fully identify all homogeneous pixels of one object, leading to their fragmentation (formation of gaps). In addition, these methods can't detect a partial stops of objects, because them are usually used along with other methods.

Methods of optical flow are based on the fact that the fragment containing moving objects can calculate the direction and magnitude of the velocity of each pixel. Information on the optical flow is used for spatial image segmentation: a group of nearby pixels which moving with the same speed can be considered a moving subject. You can also get information on the location, size and other parameters of the region. However, these algorithms are too demanding and sensitive to noise.

Methods of the background processing are defined by relative simplicity of mathematics. Characteristics of the most common of them are presented in [1-4]. The essence of these methods is compared pixel by pixel the current frame with the background model, which is the video frame without moving objects, as well as periodic updating of the model. The main drawbacks of these methods are errors in the classification of pixels at periodic background changes and delays in updating the model. Therefore, an important task is the development of methods for segmentation of moving objects, which are resistant to adverse external influences on video streams in real-time.

I. Modification of the Adaptive Segmentation Algorithm Based on Using a Mixture of Normal Distributions

This algorithm belongs to a group of algorithms which are based on probabilistic models and consider changing value of each pixels of video frame as a pixels process, ie is time series consisting of scalar values for grayscale images and vectors - for colored. At any time t for each pixel (x_0, y_0) in which are known his previous values:

$$\{X_1 \dots X_t\}=\{I(x_0, y_0, i): 1 \le i \le t\}, \quad (1)$$

where $X_i = I(x_0, y_0, i)$ – brightness of a pixel (x_0, y_0) in i-th frame.

Using a mixture of normal distributions we create pixel model of video stream, and with each new video frame - refresh this model and classify each pixel to its affiliation with the background or a moving object:

$$S \sim \sum_{p=1}^{k} (w_p \cdot N(x, \mu_p, \delta_p^2)), \quad (2)$$

where:

$$N(x, \mu_p, \delta_p^2) = \frac{1}{\sqrt{2\pi}\delta} e^{-\frac{(x-\mu)^2}{2\delta^2}}.$$

Each of the everyone summands amounts (2) corresponds to the pixel process that is characterized by the parameters of the normal distribution (expectation and variance) and the weighting factor w_p which is a measure of how often this process occurs in the video stream. Parameter k - the maximum number hausians, belongs to the interval [3; 5].

In the first video frame we initializes pixel model, assume that all pixels of the video frame are the background for each pixel we create pixel process with the following parameters:

$$w = 1, \mu = I(x_0, y_0, 1), \delta^2 = \delta_{std}^2 \quad (3)$$

where δ_{std}^2- - the variance of the normal distribution by default, accept within [50; 65].

For each video frame and the next for each pixel in it, follow these steps:

Step 1. Find the process which corresponds to the current value of the pixel brightness $I(x_0, y_0, t)$. For each pixel in the process of applying the threshold:

$$\frac{|\mu - I(x_0, y_0, i)|}{\delta} \le e \quad (4)$$

If the current pixel brightness values $I(x_0, y_0, t)$ of a process satisfies the condition (4), this process taking as current and update its statistics, or create a new process.

Step 2. Create a new process. Estimates of the expectation and variance to choose similar to model initialization pixel by the expressions (3). If the number of processes in the model is k, then search the process with the lowest weight, while the current weight w do not change, and other parameters equate to parameters of a new process, and this process is new processable and go to step 5. If the number of processes for the current pixel reached maximum value, then add another process with zero weight and go to step 4.

Step 3: Update statistics of the current process. The value of the expectation and variance updated using recursive low-pass filter smoothing. Denote, μ_{t-1}, $\delta_{t-1}^2$the parameters of the current process in the previous step, and μ_t, δ_t^2- at the current value, then:

$$\mu_t = (1-\alpha_1) \cdot \mu_{t-1} + \alpha_1 \cdot I(x_0, y_0, t), \quad (5)$$
$$\delta_t^2 = (1-\alpha_2) \cdot \delta_{t-1}^2 + \alpha_2 \cdot (\mu_t - I(x_0, y_0, t))^2,$$

where α_1, α_2- filter options that allow you to adjust the speed upgrade.

Step 4. Update the weights of all processes. Denote the $w_{i,t-1}$ weight of the i-th process in the previous step, and $w_{i,t}$ the weight of the i-th process on this, then:

$$w_{i,t} = (1-\alpha_3) \cdot w_{i,t-1} + \alpha_3 \cdot M(i) \quad (6)$$

where α_3 - parameter, which is responsible for the rate of change of weight, $M(i)$- a function that returns 1 if i equals index of the current process and 0 - when all other values i.

Step 5. Classification pixel and forming a binary mask. Classification - the weight of the current process is compared with a threshold value that takes values in the interval [0; 1]. If the weight of the current process is greater than the threshold, the pixel is classified as a background, or the pixel, owned by a moving object.

Step 6. Apply the algorithm eliminating moving shadows.

Step 7. Application algorithm of median filtering for binary mask.

Classification pixel (step 5) admits belonging to the background of several processes, because the presence of dynamic background algorithm after a certain period of time adapting to cyclical movements background. For recognition accuracy of cyclical movements, weight threshold pixel process should belong to the interval [0,3; 0,5]. Also, due to the expression (5) algorithm is able to adapt to the smoother changing light in video frame.

III. Adaptive Algorithm for Segmentation of Moving Objects Based on Selective Updated Model Background

Structural, algorithm consists of the following steps:

Step 1. Initialize the original background model. Background model $B_t(x, y)$ which is formed at time t is initialized by splitting the input image $I(x, y)$ at 32x24 block of pixels. For each block are calculated difference values of pixels in the previous and the current frame, and is determined by the weight of the block, which is equal to the number of sustainable pixels - difference in which the previous and the current frame is equal to 0. Selective background model is updating at a time when all blocks containing 95-100 % stable points.

Step 2: Update the background model by using selective temporal median filter. For each pixel of the video frame creates a cyclic buffer with size $n = k + 1$ elements, in which turn entered the brightness value of the pixels on the k-frames, and the pixel brightness values contained in the initial background model $B_t(x, y)$ created in step 1.

The value of the cyclic buffer ordered in ascending order, and is searched value, which is located in the middle of cyclic buffer. Updating the model background $B_t(x, y)$ happening by replacing its corresponding pixel on the average value of the element a cyclic buffer.

Step 3. Formation resulting binary mask. In this step occurs calculating the difference $D(x, y)=|I(x, y) - B_t(x, y)|$ current frame of video with your updated in step 2 binary mask and computation of two thresholds cyclical buffer by the expressions:

$$Th_1(i,j) = \left(\left|P_{\frac{1}{4}k} - P_{\frac{3}{4}k}\right|\right), \qquad (7)$$

$$Th_2(i,j) = (|P_1 - P_k|)$$

where k - the number of elements of the cyclic buffer, Pi - element of the cyclical buffer that is in position i.

In the resulting binary mask $M(x, y)$, the pixel difference $D(x, y)$ is classified as being owned by a moving object, if present on the mask Th_1 and meets or in the vicinity of three pixels mask Th_2, otherwise the pixel is considered as background.

Step 4. Verification of the moving objects. For segmentation of pixels that can be attributed to small movements by generated background moving objects in the resulting binary mask $M(x, y)$ are allocated by grouping by the all connected pixels and calculates their modules gradients by the expression:

$$G_t = \sqrt{\left\|\frac{\partial I_t(i,j)}{\partial(x,t)}\right\|^2 + \left\|\frac{\partial I_t(i,j)}{\partial(y,t)}\right\|^2}, \qquad (8)$$

$$\frac{\partial I_t(i,j)}{\partial(x,t)} = I_{t-\Delta t}(i-1,j) - I_t(i+1,j) \qquad (9)$$

$$\frac{\partial I_t(i,j)}{\partial(y,t)} = I_{t-\Delta t}(i,j-1) - I_t(i,j+1) \qquad (10)$$

where (9) - gradient in the horizontal direction, (10) - gradient in the vertical direction.

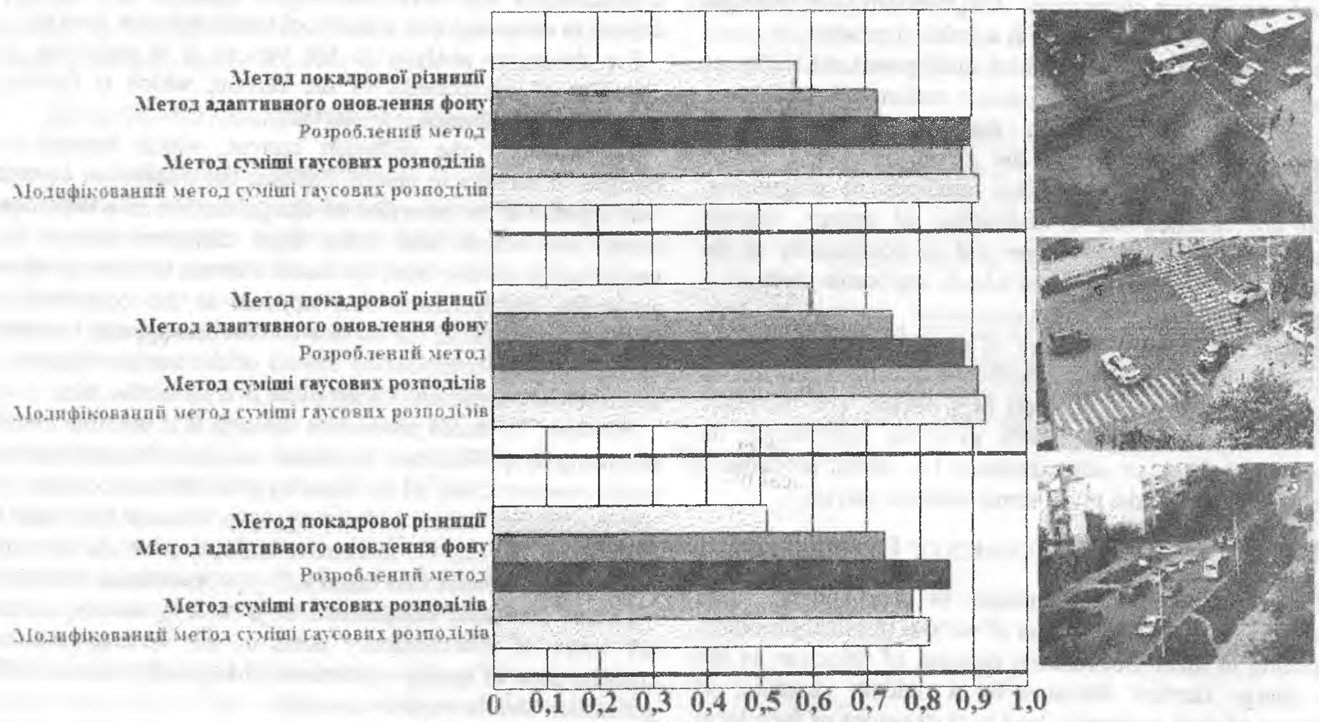
VI. Conclusion

The effectiveness of the method of segmentation is determined by the following expression:

$$Q = \frac{N_L}{N_T}, \qquad (11)$$

where N_L - number of total number of objects, N_T - number of correctly segmented objects.

The figure shows the results of comparing the effectiveness of three existing, and our methods. The results obtained for three different streams of frames to resolutions 704x576, 1080 x 960 and 640x480.



Fiq 1. Comparations of segmentation algorithm of moving objects

References

[1] Blackman, S. and Popoli, R. (1999) Design and Analysis of Modern Tracking Systems. MA: Artech House.

[2] Development of Tracking and segmentation algorithm of partial occluded moving objects: Proceedings of the 12 International Conference [«The experience of designing and application of CAD systems in microelectronics» (CADSM 2013)], (Polyana, Ukraine, February 19–23, 2013) / Lviv Polytechnik National University. – Lviv : Lviv Polytechnic, 2013. – P. 280–282.

[3] Hernandez, M., Farina, A. and Ristic, B. (2006) PCRLB for tracking in cluttered environments: measurement sequence conditioning approach, IEEE Transactions on Aerospace and Electronic Systems, 42(2), 680–704.

[4] Hwang, I., Balakrishnan, H., Roy, K. and Tomlin, C. (2004) Multiple-target tracking and identity management in clutter, with application to aircraft tracking, in Proc. 2004 American Control Conference, Boston, MA, June 30–July 2, pp. 3422–3428.