Use of stochastic adaptation in block method to estimate deformation field for image sequence

R.O. Kovalenko[1], P.V. Smirnov[2], R.M. Ibragimov[1]

[1]Ulyanovsk State Technical University, Severniy Venetz 32, Ulyanovsk, Russia, 432027
[2]"Ventra", Malaia-Semenovskaia 9, Moscow, Russia, 107023

Abstract. The article proposes a variant of the block method for estimating the deformation field for image sequence using the non-identification stochastic adaptation procedure. As a target function, the article considers the mean square inter-frame difference and the inter-frame correlation coefficient. The developed algorithm has high noise immunity and allows you to get rid of the influence of global inter-frame geometric changes. Also, by adjusting the size of the blocks, it makes it possible to remove small-sized moving objects that are not of interest (rain, snow, falling leaves, etc.), and vice versa to highlight them.

1. Introduction

Detection of the area of a moving object is a basic operation in machine vision systems because it allows you to highlight the areas of interest in the image and simplify the subsequent analysis. The task of detecting a moving object for complex cases has not yet received a final solution. The complexity of this task is caused by the possibility of various dynamic changes in the scene (smooth, sharp, local changes in light, weather changes, repetitive movement, etc.). A more complex case is observed when the background is similar to a moving object. Therefore, the development of approaches to the analysis of scene movement of image sequences, allowing in difficult conditions to effectively find areas of the image of moving objects, remains relevant.

The task of selecting the area of a moving object is usually formulated as the task of dividing the pixels of an image into two groups: the background and the foreground, where the foreground is the moving object. As the foreground can be one object or several. In both cases, the foreground objects must be detected, and if there are several objects, the moving objects must also be separated from each other.

As with many other digital image processing tasks, detection and selection of a moving object area can be implemented in both spatial and frequency area.

To detect motion in the frequency domain, most methods are based on wavelet transformations [1] and low order fractional statistics [2]. Changing the background has less effect on the result of selecting the area of the moving object by methods implemented in the frequency domain than in the methods of the spatial domain. But with this approach, there are problems with shadows [3]. Methods in the frequency domain are characterized by high computational complexity, so in practice, they are used much less often than methods of the spatial domain.

Different approaches exist to identify the area of a moving object in the spatial area. Among them, there are methods based on finding of inter-frame difference [4, 5], background subtraction [4, 6], application of statistics [5, 7], block method [8], optical flow analysis [9]. In this paper, we consider the block method.

2. Statement of the research task

Most methods for constructing a deformation field $\mathbf{H}$ are based on inter-frame image processing. In this case, the image of a moving object can be represented as some region (regions) of the investigated image having inter-frame geometric changes (IGC). Thus dividing the image into disjoint areas (blocks) and evaluating their inter-frame deformation parameters, a field $\mathbf{H}$ will be obtained. Estimation field is used to determine image blocks that correspond to a moving object, for example by threshold processing. This approach corresponds to the general principle of block methods for detecting motion [10], which are based on finding the location of blocks of the current (deformed) frame on the previous (reference) frame. To do this, the current frame $\mathbf{Z}^t$ of the image sequence is divided into many non-overlapping blocks $B_{i,j}$, where i, j are the coordinates of the centers of the blocks. The size of blocks is selected based on the size of objects whose movement needs to be detected. The solution comes down to finding the motion vector $\bar{h}_{i,j}$ of each block $B_{i,j}$ on frame $\mathbf{Z}^{t-1}$.

$$\bar{h}_{i,j} = \arg\left(\underset{v_{i,j}\in O}{\text{extremum}} \left(Q\left(i, j, v_{i,j}\right)\right)\right), \tag{1}$$

where O – is the search area, $Q\left(i, j, v_{i,j}\right)$ – is the quality function of matching blocks of the current and previous frames, which we will call the target function below. By assigning the shift $\bar{h}_{i,j}$ to the nodes of the reference grid included in block $B_{i,j}$, we obtained the deformation field $\mathbf{H} = \{\bar{h}_{i,j}\}$ for the deformed image and reference image. It became widespread, due to the fact that this approach provides rather a high efficiency at relatively low computational labor intensity [8, 10].

However, block methods assume the immobility of the background on which moving objects are detected. In practice, adjacent frames can have global mutual spatial displacements, for example, due to the movement of the camera or the structure on which the camera is mounted. Then the algorithm, based on the block method, will show the presence of motion on almost the entire frame. To overcome this disadvantage, a more complex model for determining the location of blocks $B_{i,j}$ is chosen, such as the Euclidean or similarity model [11], which includes parameters $\bar{\alpha}^{t,t-1} = (\bar{h}, \varphi, \kappa)^T$: shift in the basic axes $\bar{h} = (h_x, h_y)^T$, rotation angle φ and scale κ. The article proposes to estimate the location of blocks $B_{i,j}$ by stochastic non-identification adaptation procedure [12] to find the parameters of $\bar{\alpha}^{t,t-1}$. The algorithm is resistant to impulse noise and requires small computational costs that are virtually independent of block sizes. Block sizes are usually significantly smaller than the size of the detected object.

3. Description of the algorithm

The proposed stochastic block algorithm assumes a recurrent finding the estimates vector $\bar{\alpha}^{t,t-1}_{i,j}$ of the parameters of its position on the deformed frame for each block $B_{i,j}$ on the reference frame in accordance with the procedure [12]:

$$\hat{\bar{\alpha}}^{t,t-1}_{i,j(n)} = \hat{\bar{\alpha}}^{t,t-1}_{i,j(n-1)} - \mathbf{\Lambda}_n \bar{\beta}_n (\mathrm{J}(\hat{\bar{\alpha}}^{t,t-1}_{i,j(n-1)}, Z_n)), \tag{2}$$

where $\bar{\beta}$ – stochastic gradient of the target function $\mathrm{J}(\cdot)$; $\mathbf{\Lambda}_n$ – a positive definite diagonal gain matrix that sets the step of incrementing parameter estimates; Z_n – a local sample of image samples used to find $\bar{\beta}$ at the n iteration, $n = \overline{0, N-1}$; N – the number of iterations. Note that a local sample Z_n is independently generated for each iteration of estimation.

The algorithm was implemented for the two most common target functions: the mean square inter-frame difference (MSID) and the inter-frame correlation coefficient [13]. When using MSID for the stochastic gradient at the n-th iteration, we obtain [14]:

$$\beta_{in} = \frac{1}{2\mu\Delta x}\sum_{l=1}^{\mu}\left(\Delta\tilde{z}_x^t\left(\tilde{z}^t_{xl+\Delta x,yl} + \tilde{z}^t_{xl-\Delta x,yl} - 2z^{t-1}_{il,jl}\right)\right)\frac{\partial x}{\partial\alpha_i} + \\ + \frac{1}{2\mu\Delta y}\sum_{l=1}^{\mu}\left(\Delta\tilde{z}_y^t\left(\tilde{z}^t_{xl,yl+\Delta y} + \tilde{z}^t_{xl,yl-\Delta y} - 2z^{t-1}_{il,jl}\right)\right)\frac{\partial y}{\partial\alpha_i}, \quad (3)$$

where (x_l, y_l) – coordinates on image $\mathbf{Z}^t$ of count $\tilde{z}^{t-1}_{xl,yl}$ with coordinates (i_l, j_l) in the image $\mathbf{Z}^{t-1}$; $\tilde{z}^t_{xl,yl}$ – is the brightness of the oversampling image $\mathbf{Z}^t$ taking into account the estimates $\hat{\alpha}^{t,t-1}_{i,j(n-1)}$, obtained in the previous iteration; $\Delta x, \Delta y$ – the step of finding derivatives $\partial\tilde{z}^t_{xl,yl}/\partial x$ and $\partial\tilde{z}^t_{xl,yl}/\partial y$ using the finite difference [13], μ – local sample size Z_n. Partial derivative $\partial x/\partial\bar{\alpha}$ and $\partial y/\partial\bar{\alpha}$ are found analytically.

When using the inter-frame correlation coefficient, the expression of the stochastic gradient on the n-th iteration takes the form:

$$\beta_{in} = \frac{1}{2\mu\hat{\sigma}_{t-1}}\left[\sum_{l=1}^{\mu}\left(\frac{z^{t-1}_{il,jl} - z^{t-1}_m}{\Delta x}\left(\frac{\tilde{z}^t_{xl+\Delta x,yl}}{\sigma_{+\Delta x}} - \frac{\tilde{z}^t_{xl-\Delta x,yl}}{\sigma_{-\Delta x}}\right)\right)\frac{\partial x}{\partial\alpha_i} + \\ + \sum_{l=1}^{\mu}\left(\frac{z^{t-1}_{il,jl} - z^{t-1}_m}{\Delta y}\left(\frac{\tilde{z}^t_{xl,yl+\Delta y}}{\sigma_{+\Delta y}} - \frac{\tilde{z}^t_{xl,yl-\Delta y}}{\sigma_{-\Delta y}}\right)\right)\frac{\partial y}{\partial\alpha_i}\right], \quad (4)$$

where $\sigma^2_{\pm\Delta x} = (\mu-1)^{-1}\left(\sum_{l=1}^{\mu}\left(\tilde{z}^t_{xl\pm\Delta x,yl}\right)^2 - \mu\left(\tilde{z}^t_{\pm\Delta xm}\right)^2\right)$, $\hat{\sigma}^2_{t-1} = (\mu-1)^{-1}\sum_{l=1}^{\mu}\left(z^{t-1}_{il,jl} - z^{t-1}_m\right)^2$; $\tilde{z}^t_{\pm\Delta xm}$ and $z^{t-1}_m = \mu^{-1}\sum_{l=1}^{\mu} z^{t-1}_{il,jl}$ – the mean values of $\tilde{z}^t_{xl\pm\Delta x,yl}$ and $z^{t-1}_{il,jl}$.

The algorithm based on MSID requires less computational costs and can work already with the volume of the local sample $\mu = 1$, which allows it to be implemented in pixel-by-pixel processing. Therefore, in the proposed approach, the choice of MSID as the main target function is appropriate.

The operation of the algorithm can be simplistically described as follows. For neighboring frames that do not have mutual global IGC, their estimates of the parameters of blocks without motion will remain close to zero in contrast to blocks with motion, whose parameter estimates will converge to some non-zero values [15] (for a scale factor of 1). Such convergence is a criterion for assigning a block to motion. If neighboring frames have mutual global IGC, then the estimates of the deformation parameters of all blocks will be different from zero. In this case, the blocks corresponding to the moving object will form compact clusters. Blocks with global deformations are located on average throughout the frame, which is used as a criterion for determining global deformations [16]. The deformation parameters of moving objects are determined by subtracting the global [17] deformations.

4. Some experimental results

Figure 1 shows an example of two adjacent frames of the image sequence $\mathbf{Z}^t$, which was obtained with a microscope at a magnification of 400 times. On this figure, you can see two unicellular Sonderia organisms. An organism that is completely in the frame is in motion. Motion parameters can be written by using the similarity model: $\bar{h} = (3.6, -2.1)^T$, $\varphi = 3^\circ$, $\kappa = 1$. And the second organism is almost motionless. At the same time frames have global IGC with parameters: $\bar{h} = (1, -2.2)^T$, $\varphi = -1^\circ$ $\kappa = 1.01$. Also for a complex case, unbiased additive Gaussian noise with a signal/noise ratio of 14 dB was added to the images.

Figure 2 shows the comparative results of the inter-frame difference algorithms (Figure 2a), background subtraction (Figure 2b) and the proposed stochastic block algorithm (Figure 2c). For ease of comparison, an organism contour has been added to each image.

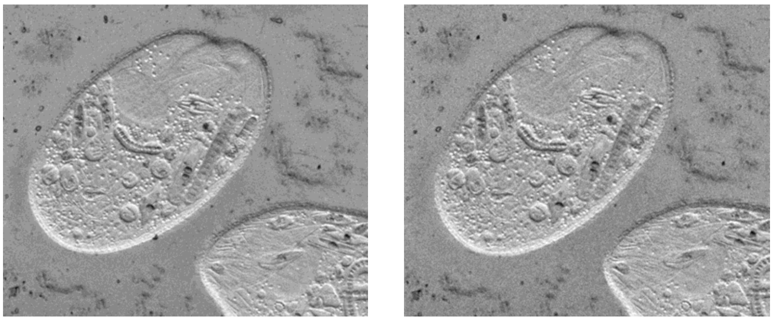

Figure 1. An example of adjacent image sequence frames.

Figure 2. The result of motion detection by different algorithms.

Figure 2 shows that the inter-frame difference and background subtraction algorithms define the second organism in motion, due to global geometric changes in adjacent frames. These two algorithms detect an area of a moving object with a large number of gaps, especially in low-contrast places where there is a small gradient of image brightness. The proposed stochastic block algorithm allocates a region of motion with almost no gaps. The gaps can only correspond to blocks in which most of the pixels relate to the background and only some of them relate to a moving object.

5. Conclusion

The developed algorithm, based on non-identification stochastic adaptation, has high noise immunity and allows us to get rid of the influence of global IGC, as well as to remove small moving objects that are not of interest. For example, in this paper, such objects were small organisms and particles, in other situations - rain, snow, falling leaves, etc. The allocation of small objects is realized by reducing the size of blocks, up to one pixel [18].

6. Acknowledgments

The reported study was funded by RFBR and Government of Ulyanovsk Region according to the research projects № 18-41-730011 and 19-29-09048.

7. References

[1] Antic, B. Efficient wavelet based detection of moving objects / B. Antic, V. Crnojevic, D. Culibrk // 16th International Conference on Digital Signal Processing, 2009. – P. 1-6.

[2] Bagci, M. Moving object detection using adaptive subband decomposition and fractional lower statistics in video sequences / M. Bagci, Y. Yardimci, A.E. Cetin // Signal Process. International Journal of Signal Processing. – 2002. – Vol. 82(12). – P. 1942-1947. DOI: 10.1016/S0165-1684(02)00321-3.

[3] Töreyin, B.U. Moving Object Detection in Wavelet Compressed Video / B.U. Töreyin, A. Enis Çetin, A. Aksay, M.B. Akhan // Signal Processing: Image Communication. – 2005. – Vol. 20(3). – P. 255-264. DOI: 10.1016/j.image.2004.12.002.

[4] Elhabian, Sh.Y. Moving Object Detection in Spatial Domain using Background Removal Techniques / Sh.Y. Elhabian, Kh.M. El-Sayed, S.H. Ahmed // Recent Patents on Computer Science. – 2008. – Vol. 1(1). – P. 32-54.

[5] Karasulu, B. Performance Evaluation Software: Moving Object Detection and Tracking in Videos / B. Karasulu, S. Korukoglu – Berlin: Springer Science & Business Media, – 2013. – 76 p.

[6] Wang, L. Extraction of Moving objects from their Background based on mulitple adaptive threshold and boundary evaluation / L. Wang, N.H.C. Yung // IEEE Trans. Intelligent transportation systems. – 2009. – Vol. 11(1). – P. 40-51. DOI: 10.1109/TITS.2009.2026674.

[7] Kutsov, R.V. Algorithms for detecting a moving object in an image / R.V. Kutsov, A.P. Trifonov // Izvestiya RAN. Theory and control systems. – 2006. – Vol. 3. – P. 129-138.

[8] Grishin, S.V. A review of block-based methods for estimating motion in digital video signals / S.V. Grishin, D.S. Vatolin, A.S. Lukin, S.Iu. Putilin, K.N. Strelnikov // Software systems and tools: Thematic collection. – 2008. – Vol. 9. – P. 50-62.

[9] Zolotykh, N.Iu. An overview of the methods for searching and tracking vehicles on the video stream / N.Iu. Zolotykh, V.D. Kustikova, I.B. Meerov // Vestnik of the Nizhny Novgorod University. N.I. Lobachevsky. – 2012. – Vol. 5(2). – P. 348-358.

[10] Zaqout, I.S. An efficient block-based algorithm for hair removal in dermoscopic images / I.S. Zaqout // Computer Optics. – 2017. – Vol. 41(4). – P. 521-527. DOI: 10.18287/2412-6179-2017-41-4-521-527.

[11] Pons, D. Computer vision. Modern approach / D. Pons, Zh. Forsait – Moscow: Viliams, 2004. – 926 p.

[12] Tashlinskii, A.G. Estimation of spatial deformation parameters of image sequences / A.G. Tashlinskii – Ulyanovsk: ULSTU, 2000. – 132 p.

[13] Tashlinskii, A.G. Methods of finding gradient estimates of target function for measurement of images parameters / A.G. Tashlinskii, P.V. Smirnov, L.Sh. Biktimirov // Pattern recognition and image analysis. – 2011. – Vol. 21(2). – P. 339-342. DOI: 10.1134/S1054661811021057.

[14] Tashlinskii, A.G. Analysis of methods of estimating objective function gradient during recurrent measurements of image parameters / A.G. Tashlinskii, P.V. Smirnov, S.S. Zhukov // Pattern recognition and image analysis. – 2012. – Vol. 22(3). – P. 399-405. DOI: 10.1134/ S1054661812020186.

[15] Tashlinskii, A.G. Criteria for stopping the process of pseudo-gradient image binding based on the analysis of convergence of estimates of binding parameters / A.G. Tashlinskii, L.Sh. Biktimirov, P.V. Smirnov // High technology. – 2013. – Vol. 14(5). – P. 22-25.

[16] Tashlinskii, A.G. A way to predict parameters of image registration by estimating inter-frame deformation of local fragments / A.G. Tashlinskii, S.V. Voronov, P.V. Smirnov // Pattern recognition and image analysis. – 2014. – Vol. 24(1). – P. 179-184. DOI: 10.1134/ S1054661814010192.

[17] Tashlinskii, A.G. Pseudo-gradient estimation of spatial deformations of an image sequence / A.G. Tashlinskii // High technology. – 2002. – Vol. 3. – P. 32-43.

[18] Tashlinskii, A.G. Per-pixel assessment of scene motion by video sequence / A.G. Tashlinskii, P.V. Smirnov, M.G. Tsaryov // Automation of management processes. – 2017. – Vol. 4(50). – P. 67-74.