System for detecting and tracking moving objects

A. Tarasov
Ryazan State Radio Engineering University named after V.F. Utkin
Russia, Ryazan
vb2005@yandex.ru

M. Nikiforov
Ryazan State Radio Engineering University named after V.F. Utkin
Russia, Ryazan
nikiforov.m.b@evm.rsreu.ru

***Abstract*— This paper considers the construction of a system for detecting and tracking moving objects. It is proposed to pre-process the frame using digital image stabilization algorithms based on optical flow. To detect objects, it is supposed to use the longest optical flow vectors formed after stabilization, and to implement tracking using several classical algorithms using a prefetch mechanism built on classification neural networks.**

Keywords*— *Optical flow, object tracking, technical vision

I. INTRODUCTION

Tracking objects is the process of determining the location of some area of interest in the frame in real time. The area of interest is understood as a certain rectangular fragment in the image, in which the target of interest to the user is located.

The object of scientific research is the task of searching for universal (that is, the system is not pre-trained to detect specific types) objects in the video stream. In this regard, the subject of research is the methods of digital image processing used to solve problems of searching for an object, its automatic detection and tracking of an object in a video stream.

The purpose of this work is to develop methods and algorithms for detecting and tracking moving objects that are invariant to projective transformations and shooting conditions and to implement software tools for detecting and tracking.

II. PRELIMINARY PROCESSING

In the situation that the camera is stationary, one of the approaches to object detection can be the subtraction of the background component. To do this, formed a new image based on two adjacent frames:

$$I'(x,y,t) = I(x,y,t) - I(x,y,t-1), \quad (1)$$

I(x, y, t) – brightness intensity in *x* and *y* coordinates and at time *t*.

However, in most cases, the camera or the background is moving, therefore, it is worth calculating the difference with an offset:

$$I'(x,y,t) = I\left(x+u_x, y+u_y, t\right) - I(x,y,t-1), \quad (2)$$

u – frame offset on *x* & *y* axis.

The offset of the video system has a negative impact on the accuracy of detection and tracking of objects. In this work, optical flow stabilization is used for the following reasons:

- Sparse optical flow is calculated in a short time (less, than 10 ms.)
- Optical flow maps have a high degree of relevance. Even having information about the movement of several vectors, one can draw conclusions about the direction of movement;
- The optical flow allows you to detect not only linear movement, but also rotation, tilt and distance / approximation of the video system;
- After normalization, optical flow maps can be used both for detecting moving objects and for tracking.
- In this paper, the determination of the movement of the video system is performed on the basis of the optical flow of Lucas-Kanade [1].

More details about this method are described in our work [3].

III. DETECTION OF MOVING OBJECTS

The object detection algorithm includes: selection of a moving object in which the operator is interested in tracking; identification of moving objects; detection and selection of objects of interest, analysis of the velocity vectors of selected objects, localization with preservation of object coordinates and their characteristics. To increase the probability of automatic detection of moving objects, a combination of methods is used (according to the diagram in Figure 1)

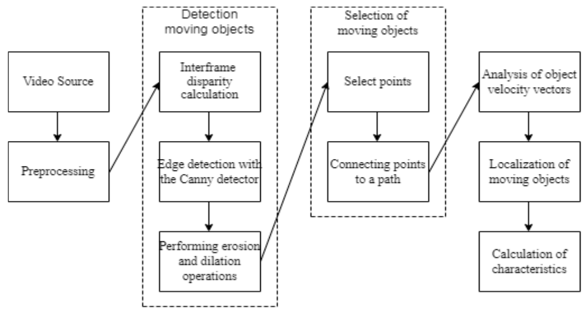

Fig. 1. A system for detecting and highlighting moving objects scheme

The choice of an object for tracking can be carried out by the operator, who selects the area of interest, or in automatic mode. The region of interest is a rectangular fragment of the image, which completely contains the desired object, while the size of the fragment should be as small as possible. The operator selects an area with an object. In subsequent frames, a fragment is searched for, identical in size, having the smallest difference with the object of interest.

A. *The moving areas selection*

The selection of moving objects begins with the calculation of interframe disparity and the division of all image points into two classes (background and moving points), by binarization (Figure 2).

Fig. 2. Original image (school lesson)

B. *Grouping moving pixels into objects*

To solve the grouping problem, a number of approaches based on the texture difference, as well as on the basis of optical flow vectors, were considered [2]. It should be noted that these approaches lose in accuracy to distance-based region division algorithms, however, they can be used to separate connected objects (as, for example, in Figure 3).

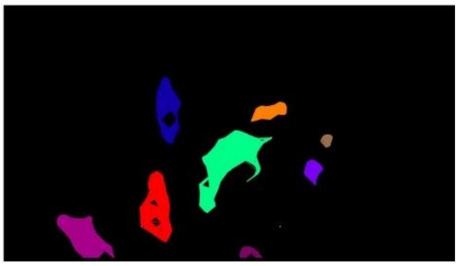

Fig. 3. Moving areas

IV. TARGET TRACKING

To track objects in this paper, it is proposed to use a prefetching mechanism based on the preliminary selection of such a tracking algorithm that, under given conditions, would demonstrate the best results in terms of accuracy.

It is proposed to use convolutional classification neural networks as a prefetching mechanism. The interframe disparity in the area of the tracked object is fed to the input of such a network, and the identifier of the most optimal method for the given conditions is given at the output.

A. *Tracking algorithms*

- **Template comparison algorithm (CS)**
- **The fast comparison algorithm (FCS)**
- **Algorithm based on the calculation of the optical flow (OF)**.
- **Algorithm based on the matching of key points (KP)**
- **Linear regression algorithm with Kalman filter (LR)**

B. *Algorithms comparsion method*

There are several approaches to determining the error of the tracking algorithm. It can be evaluated as well as work in a separate frame, and throughout the plot.

Most sources [4] are based on the calculation of the "intersection over union" indicator, which is defined as the ratio of the area of the figure formed by the intersection of the set from the predicted location of the object and the set from the reference (marked) location of the object to the area of the union of these sets.

Below are the results of the algorithms described above on the video sequences presented as part of the Visual Object Tracking 2014 contest (Figure 4).

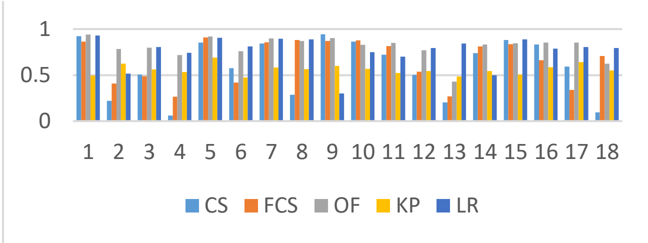

Fig. 4. Mean IoU metric chart for 18 scenarios for 5 different algorithms

C. *Results of work for prefetching mechanism*

Table 1 provides information on the operation of the prefetch mechanism. For each of the presented methods, the accuracy of work is described, the highest value of accuracy is determined and the value when using the prefetch mechanism

TABLE I. ACCURACY FOR ALGORITHMS

<table>
<tr><th colspan="6">Accuracy for algorithms</th></tr>
<tr><td>CS</td><td>FCS</td><td colspan="2">OF</td><td>KP</td><td>LR</td></tr>
<tr><td>0.60</td><td>0.52</td><td colspan="2">0.60</td><td>0.43</td><td>0.49</td></tr>
<tr><th colspan="6">Prefetching algorithm</th></tr>
<tr><td colspan="2">Best matching value</td><td colspan="2">Prefetching algorithm</td><td colspan="2">Time cost of prefeching algorithm</td></tr>
<tr><td colspan="2">0.75</td><td colspan="2">0.68</td><td colspan="2">5 ms</td></tr>
</table>

V. CONCLUSION

The methods and approaches described above can be used in the construction of tracking systems for moving objects. The elimination of the movement of the optical system reduces the error probability by 40%, the mechanism for selecting moving objects significantly outperforms the methods described in the works due to the stabilization and classification and linking of vectors. The maintenance method prefetch mechanism reduces the error by 15-20% by choosing the most efficient method. Despite the large number of approaches used, this complex fits into the real-time format (24 frames per second) on modern processors.

REFERENCES

[1] Lucas, B. D. An iterative image registration technique with an application to stereo vision / B. D. Lucas, T. Kanade // Proceedings of Imaging Understanding Workshop. – 1981. – P. 121–130

[2] Nikiforov, M. B. System for automatic detection of moving objects in a video sequence / M. B. Nikiforov, A. S. Tarasov // XXI century: results of the past and problems of the present plus. – 2022. – Vol. 11. – P. 39–44

[3] Tarasov, A. Intelligent video computer tracking system / A. Tarasov, Y. Muratov, M. Nikiforov, A. Baranchikov // 13th International Conference ELEKTRO. – 2020. – Vol.13. – P. 913.

[4] Loey, M. Fighting against COVID–19: A novel deep learning model based on YOLO–v2 with ResNet–50 for medical face mask detection / M. Loey, G. Manogaran, M. H. N. Taha, N. E. M. Khalifa // Sustainable Cities and Society. – 2021. – Vol. 65. – P. 1026