

STUDY OF DETECTION AND TRACKING ALGORITHMS OF MOVING OBJECTS IN VIDEO SEQUENCES FROM VIDEO SURVEILLANCE CAMERAS

Anatolii Babaryka

INTRODUCTION

The wide availability of video surveillance systems has led to a situation where we have the opportunity to place a sufficiently large number of surveillance cameras on controlled objects, but at the same time there is a problem of monitoring the situation from them. Video surveillance operators are physically unable to simultaneously monitor information from 10-15 or even more video channels. That is why, a logical step in the development of video surveillance systems is the introduction of video analytics functions.

One of the detectors that are used in modern intelligent video surveillance systems is the conditional line crossing detector. The logic of this detector is that the operator in the software interface builds a conditional line, at the intersection of which certain objects (persons, vehicles or any moving objects) there is a visual (and / or sound) alarm.

Therefore, the introduction of this type of detectors in video surveillance systems, which are deployed at facilities with perimeter security, will effectively monitor the situation in the sectors of video surveillance cameras on a certain number of video channels at the same time.

Survey of algorithms for detection and tracking of objects in video sequences was conducted in the works of a wide range of scientists, such as Boris Babenko, P. Viola and M. J. Jones, Zhengxia Zou, Zhenwei Shi, Yuhong Guo, Redmon J., Liu W., Anguelov D., Erhan D., Szegedy C., S. Reed, C.-Y. Fu, A. C. Berg, et al. In the works of these authors, considerable attention is paid to the analysis, development and development of these algorithms.

For the purpose of further practical implementation of the problem of detecting the intersection of the conditional line by moving objects, it is advisable to analyze the existing methods of detecting and tracking moving objects.

1. Classification of maintenance algorithms

In modern video surveillance systems with functions of automated processing of video information, one of the important tasks is the detection

of moving objects, construction of trajectories of their movement and analysis. Difficulties in the implementation of the task of tracking detected objects due to changes in the environment, movement of the object and changing its shape during movement, interference, etc.

In article¹, based on the analysis of the current state of functioning of the departmental surveillance subsystems included in the integrated information-telecommunication system «Gart» of the State border guard service of Ukraine, was identified a list of tasks that can be performed by the specified subsystem with the introduction of intelligent video processing (video analytics).

One of such tasks of intelligent video surveillance systems is detection of video sequences of crossing objects of a conditional line. To implement the above task, it is necessary to solve the following particular problems, namely: identification of objects, their localization, tracking from frame to frame and fixing the moment of intersection of a group of pixels, to connect the «object of interest» with pixels that belong to the «conditional line».

Based on certain tasks, the actual task is to conduct a study of modern algorithms for detection, tracking and analysis of the features of the trajectory of moving objects in the sectors of surveillance cameras.

Let us define a list of restrictions and assumptions that characterize the features of the functioning of the mentioned systems in video surveillance systems installed at checkpoints across the state border and on the territory of military camps, where the «objects of interest» are vehicles and people; the speed of detection of a moving object is not critical (is a few seconds); the load on the hardware must be minimized; the algorithm must be guaranteed to detect moving objects.

Taking into account certain limitations and assumptions, we will analyze the existing algorithms for detecting, tracking and analyzing the features of the trajectory of moving objects. Under the concept of object detection we will understand the process of determining the location of the object in the frame.

Object detection detectors are usually more computationally intensive and, accordingly, have lower performance than object tracking algorithms.

Therefore, for optimal use of resources of electronic computer technology are used hybrid approaches, which are built on the concept of

¹ Katerynychuk I., Babaryka A. (2018) Doslidzhennya algoritmiv viyavleniya ta suprovodzhennya ruhomih ob'ektiv u videoposlidovnostyah z kamer videosposterezheniya. [Analysis of the Technologies of the Functioning Specific Video Surveillance Systems and the Definition of their Directions of Improvement]. *Collection of scientific works of the National Academy of the State Border Guard Service of Ukraine. Series: Military and Technical Science.* no. 3(77). pp. 246–259 (in Ukrainian).

combining algorithms of detection and tracking of objects in one algorithm. The work of such algorithms is divided into two stages:

1. The detection stage, during which the algorithm for detecting new objects and objects that were «lost» by the tracking algorithm during the other stage. For each detected object, the object tracker is added or updated with new bounding box coordinates. This stage lasts a certain number of frames, and each time repeats.

2. The tracking phase begins when the detection phase stops.

At this point, algorithm tracks the object as it moves in the frame using the created object tracker. This procedure continues for a certain number of frames until the discovery phase begins again, after which the follow-up phase begins again.

To support objects in the video, it is necessary to solve the problem of detection of such objects. The solution of this problem is assigned to the object designers. Historically, the process of development of algorithms of object detection can be divided into two periods:

- the period of identification of objects by traditional methods – until 2014;
- the period of object detection based on deep learning methods after 2014². (Fig. 1)

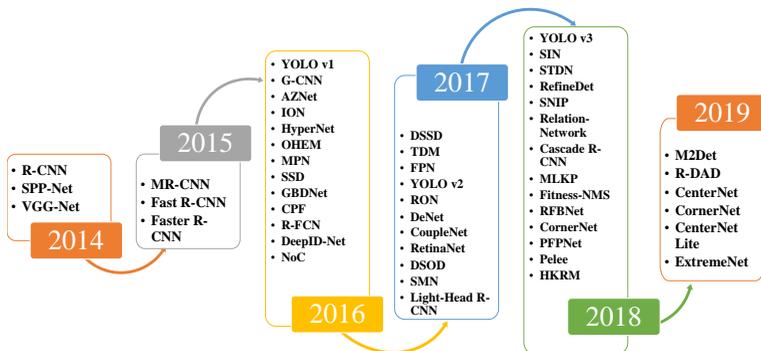


Fig. 1. Timeline of development of object detection algorithms from 2014 to 2019

²Zou Z., Shi Z., Guo Y., & Ye J. (2019) Object Detection in 20 Years: A Survey. ArXiv, abs/1905.05055. (in English)

Until 2014, most research efforts have focused on identifying a specific category of objects, such as pedestrians, faces, vehicles, etc., using pre-created sets of relevant features.

The easiest way to solve the object detection task is to reduce it to a classification task. To identify an object, you must take a specific classifier and apply it to areas of the image. To play such areas use the method of «sliding window».

So, in 2001, P. Viola and M. Jones proposed a method that allowed real-time detection of faces in images^{3,4}. The Viola-Jones detector uses a sliding window method that allows scaling and sliding window positions to detect if there are pixel sets in any window that match the face image. Haar cascades are used as a classifier (Fig. 2).

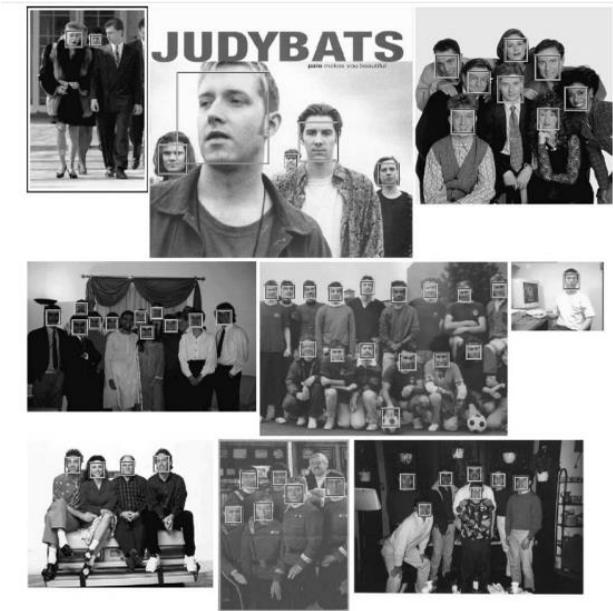


Fig. 2. The result of the Viola-Jones algorithm⁴

³ Viola P. A., & Jones M. J. (2001) Rapid object detection using a boosted cascade of simple features. *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001, 1*, 1-1. DOI: 10.1109/CVPR.2001.990517. (in English)

⁴ Viola P. A., Jones M. J. (2001) Robust real-time face detection. *Proceedings Eighth IEEE International Conference on Computer Vision. ICCV 2001, no. 2*, pp. 747-747. DOI: 10.1023/B:VISI.0000013087.49260.fb. (in English)

In 2005, N. Dalal and B. Triggs proposed a method called the HOG (Histogram of Oriented Gradients)⁵. The HOG detector scales the input image several times, keeping the size of the detection window unchanged. The basic idea of the algorithm is the assumption that the appearance and shape of the object in the image can be described by the intensity gradient distribution function.

In 2008 Felzenszwalb, P.F., McAllester, D.A., & Ramanan, D proposed a method called the DPM⁶.

In 2012, at ImageNET, a surefire win is the algorithm based on the use of convolutional neural networks (AlexNet) and this becomes a turning point in the development of neural networks.

State of the art object detection algorithms can be divided into the following categories:

- algorithms are based on «two-stage detection»;
- algorithms are based on «one-stage detection»;
- algorithms are based on point-based detection;
- algorithms are based on other author’s approaches.

In two-stage detection, the algorithm in the first stage selects the regions in which the objects of interest are likely to be located, and in the second stage the classification problem is solved. R-CNN was the first two-stagemethod. After it, appeared such detectors as Fast-RCNN, Faster R-CNN, Mask R-CNN and others. Two-stage architectures have high accuracy, but for practical application, frame processing speed very important.

One-stage algorithms solve the detection problem as a regression task. Such detectors are a single convolutional neural network, which outputs localized regions of objects of interest, their classification and the correction of the bounding box. You Only Look Once – YOLO⁷ was the first one-stage method (Fig. 3).

⁵ Dalal N., Triggs B. (2005) Histograms of oriented gradients for human detection. *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '05)*, 1, 886-893 vol. 1. (in English)

⁶ Felzenszwalb P. F., McAllester D. A., Ramanan D. (2008) A discriminatively trained, multiscale, deformable part model. *2008 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1–8. DOI: 10.1109/CVPR.2008.4587597. (in English)

⁷ Redmon J., Divvala S.K., Girshick R.B., Farhadi A. (2015) You Only Look Once: Unified, Real-Time Object Detection. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 779–788. DOI: 10.1109/CVPR.2016.91. (in English)

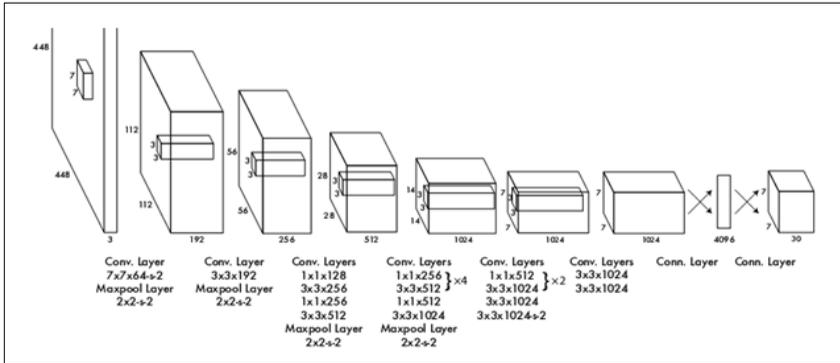


Fig. 3. YOLO architecture⁷

Further, based on the ideas proposed in the YOLO, was developed SSD (Single Shot Multi Box Detector)⁸, RetinaNet⁹ and others.

In 2019, researchers proposed a new approach to object detection that was implemented in CenterNet algorithms¹⁰.

Under the concept of tracking detected objects, we will understand the process of localization of a moving object or several moving objects on a sequence of frames.

Tracking algorithms, in contrast to detection algorithms, are able to solve the problems of such adverse conditions as occlusion (the object is partially or completely blocked), motion blur, complex background, lighting changes.

Maintenance algorithms can be classified according to the following criteria:

1. By the number of objects accompanied by:

- for single object tracking-SOT (single object tracker);
- for tracking multiple objects-MOT (multiple object tracker)

2. According to the method of support:

- tracking on the basis of detection (frames with previously identified object, are consistently processed by the tracker, which foresees the location of the object on the next frame and forms a trajectory);

⁸ Liu W., Anguelov D., Erhan D., Szegedy C., Reed S.E., Fu C., Berg A.C. (2016) SSD: Single Shot MultiBox Detector. *ECCV*. DOI: 10.1007/978-3-319-46448-0_2. (in English)

⁹ Lin T., Goyal P., Girshick R.B., He K., Dollár P. (2017) Focal Loss for Dense Object Detection. *2017 IEEE International Conference on Computer Vision (ICCV)*, pp. 2999–3007. DOI: 10.1109/ICCV.2017.324. (in English)

¹⁰ Zhou X., Wang D., Krähenbühl P. (2019) Objects as Points. *ArXiv, abs/1904.07850*. (in English)

– maintenance without detection (manual initialization of the «object of interest» on the first frame is required).

3. The method of obtaining frames:

– support from the video stream recorded on the information screen (the algorithm for analysis can use not only the previous frames, but also the next ones);

– video stream tracking in real-time mode (the algorithm can use only previous frames for analysis).

4. According to the method of organization of training strategy:

– real-time learning (algorithm analyzes pre-initialized frame and multiple previous frames)

– training on the basis of previously received information about the «object of interest».

One group of scientists divides the existing methods of tracking identified objects into the following categories: methods are based on tracking by «control» points, methods are based on the centers of mass of objects (core), methods are based on tracking along the contours of objects^{11,12,13}.

Another group of scientists divides the methods of maintenance into categories:

– tracking by the sites of objects;

– tracking by the contours;

– tracking by the boundaries^{14,15,16}.

In our research, we tend to the opinion of the first group of scientists on the classification of tracking methods into such groups as the methods are based on tracking by «control» points, the goals are based on tracking by the

¹¹ Patil H., Bhagat K. (2015) Detection and tracking of moving object: A survey. *Int. Journal of Engineering Research and Applications*. vol.5 no. 11. pp.138–142. (in English)

¹² Grandham Sindhuja, Renuka Devi. (2015) A survey on detection and tracking of objects in video sequence. *International Journal of Engineering Research and General Science*. vol. 3 no. 2. (in English)

¹³ Yilmaz A., Javed O. A., Shah M. (2006) Object tracking: A survey. *ACM Comput. Surv.*, vol. 38, no. 13. (in English)

¹⁴ Barga Deori and Dalton Meitei Thounaojam. (2014) A survey on moving object tracking in video. *International Journal on Information Theory (IJIT)*. vol.3, no.3 pp. 31–46. DOI: 10.5121/ijit.2014.3304. (in English)

¹⁵ Ann Maria Jacob and J Anitha (2012) Inspection of various object tracking techniques. *International Journal of Engineering and Innovative Technology*. vol. 2, no. 6, pp. 118–124. (in English)

¹⁶ Ojha S., Sakhare S. (2015) Image processing techniques for object tracking in video surveillance – A survey. *2015 International Conference on Pervasive Computing (ICPC)*, pp. 1-6. DOI: 10.1109/PERVASIVE.2015.7087180. (in English)

centers of mass of objects (core), the methods are based on tracking by the contours of objects.

2. Analysis of modern tracking algorithms

We chose Python 3.6 and OpenCV as our development environment. The OpenCV library since version 3.4 contains such algorithms of maintenance of objects as: Boosting, MIL, KCF, TLD, MedianFlow, GoTurn, MOSSE, CSRT.

For experimental research, most often used sets such as Object Tracking Benchmark, Visual Object Tracking, etc. For example, to assess the results of these algorithms was taken video consequence from video surveillance cameras that operate in conditions close to the conditions of the real application of video surveillance systems of the State guard border service of Ukraine. For the research we chose the following algorithms:

The Boosting method is based on real-time object tracking and is based on the AdaBoost algorithm. This method was proposed in article¹⁷.

The CSRT method is described in article¹⁸, and is based on the concept of discriminatory correlation filters (DCF). Despite the use of simple functions such as HoG and Colornames, this method works on a par with trackers, which are built on more computationally complex neural network methods and provides information processing in real time.

TLD was presented in article¹⁹ is based on the concept of splitting the long-term maintenance task into short-term maintenance, training and identification. The tracker tracks the object from frame to frame, the detector localizes all previously detected objects and corrects the tracker as needed. The learning process evaluates the detector errors and updates it to avoid these errors in the future. It is based on the MedianFlow median flow tracking algorithm. It is characterized by being able to handle rapid movements, partial occlusions, the absence of objects and the like.

MedianFlow is a method described in article²⁰. In this paper proposed a measure, Forward-Backward error, that estimates reliability of a trajectory.

¹⁷ Grabner H., Grabner M., Bischof H. (2006) Real-Time Tracking via On-line Boosting. *BMC*. (in English)

¹⁸ Lukezic A., Vojir T., Zajc L.C., Matas J.E., Kristan M. (2017) Discriminative Correlation Filter Tracker with Channel and Spatial Reliability. *International Journal of Computer Vision*, vol. 126, pp. 671–688. DOI: 10.1109/CVPR.2017.515. (in English)

¹⁹ Kalaz Z. (2011) Tracking-Learning-Detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, pp. 1409–1422. DOI: 10.1109/TPAMI.2011.239. (in English)

²⁰ Kalal Z., Mikolajczyk K., Matas J.E. (2010) Forward-Backward Error: Automatic Detection of Tracking Failures. *2010 20th International Conference on Pattern Recognition*, pp. 2756–2759. (in English)

A validation trajectory is constructed by backward tracking and compared to the trajectory in question. The implementation only involves applying the same tracking algorithm on a reversed sequence of images. Suitable for smooth and predictable movements of objects that fit completely into the frame image²⁰.

The algorithm KCF (Kernelized Correlation Filter) was presented in articles^{21,22} uses the properties of the circulant matrix to increase the speed of information processing. The methods of circulants and regularization by Tikhonov are used for calculations in the cortical filter. The algorithm is based on the use of the directed gradient histogram (HOG) method.

The MOSSE (Minimum Output Sum of Squared Error) algorithm was presented in article²³. Works with grayscale images and is based on the use of adaptive correlation filters to accompany visual objects.

The MIL was presented in article²⁴. This paper presents a new type of correlation filter, a Minimum Output Sum of Squared Error (MOSSE) filter, which produces stable correlation filters when initialized using a single frame. A tracker based upon MOSSE filters is robust to variations in lighting, scale, pose, and non-rigid deformations while operating at 669 frames per second. Occlusion is detected based upon the peak ratio, which enables the tracker to pause and resume where it left off when the object reappears. The method trains the classifier in real-time separation of the object from the background.

MeanShift is the method proposed in article²⁵. It is based on the technique of analyzing the feature space for the order of the maximum probability density. The MeanShift algorithm tries to find the area of the video frame that is locally most similar to the previously initialized area.

²¹ Henriques J., Caseiro R., Martins P., Batista J. (2012) Exploiting the Circulant Structure of Tracking-by-Detection with Kernels. *ECCV*. DOI: 10.1007/978-3-642-33765-9_50. (in English)

²² Henriques J.F., Caseiro R., Martins P., Batista J. (2015) High-Speed Tracking with Kernelized Correlation Filters. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.37, pp. 583–596. DOI: 10.1109/TPAMI.2014.2345390. (in English)

²³ Bolme D. S., Beveridge J. R., Draper B. A., Lui Y. M. (2010) Visual object tracking using adaptive correlation filters. *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 2544–2550. DOI: 10.1109/CVPR.2010.5539960. (in English)

²⁴ Babenko B., Yang M., Belongie S. (2009) Visual tracking with online Multiple Instance Learning. *2009 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 983–990. DOI: 10.1109/CVPR.2009.5206737. (in English)

²⁵ Comaniciu D., Meer P. (2002) Mean Shift: A Robust Approach Toward Feature Space Analysis. *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, pp. 603–619. DOI: 10.1109/34.1000236. (in English)

CamShift is a method proposed in²⁶ based on the segmentation algorithm proposed by Gary Bradsky in 1998.

GOTURN is a method proposed in article²⁷ based on the use of convolutional neural map. Due to this, the method without any additional settings in real time solves the problem of object maintenance. The specified method on a way of the organization of strategy of training belongs to algorithms working on the basis of previously received information on object of interest, that is in the course of work you do not study, and uses already ready data models. The disadvantage of this method is that it does not handle occlusions, but at the same time it is reliable enough to accompany the object in complex conditions such as changes in lighting and deformation.

3. Eperimental results and considering the analysis of the results of VOT

On the basis of the experiment, taking into account the pre-defined tasks that should implement intelligent video surveillance systems of the State border guard service, we have reached the following conclusions:

1. To implement the function of detecting the intersection of objects crossing the conditional line (detection of objects approaching or crossing the fence), it is more expedient to use algorithms based on neural network methods, since a person who approaches the fence during his movement can crouch, lie down, bend and group while changing the shape of his visual display. During the experiment, the algorithms are based on the use of Haar cascades and histograms of directional gradients (HOG) when changing the geometric shape of the detected moving object, there is no intersection of the conditional line, while reducing the threshold value of the area of the detected moving object leads to erroneous results of the algorithm due to external factors (movement of tree crowns, wind gusts, changes in lighting, etc.).

2. To implement the function of counting the number of persons and/or vehicles that cross a conditional line defined by the operator and do not significantly change their geometric proportions during movement, acceptable results of the work were shown by the algorithms KCF, MIL, MOSSE, GoTurn (neural networks algorithm). Unsatisfactory results were shown by the algorithms CSRT, TLD, Boosting, MedianFlow, MeanShift, Cam-Shift (delays in processing video frames, detection of moving objects).

²⁶ Bradski G. R. (1998) Real time face and object tracking as a component of a perceptual user interface. *Proceedings Fourth IEEE Workshop on Applications of Computer Vision. WACV'98 (Cat. No.98EX201)*, pp. 214-219. DOI: 10.1109/ACV.1998.732882. (in English)

²⁷ Held D., Thrun S., Savarese S. (2016) Learning to Track at 100 FPS with Deep Regression Networks. *ECCV*. DOI: 10.1007/978-3-319-46448-0_45. (in English)

Also, on the basis of the analysis of the results of VOT-2018²⁸ (table 1), it was determined that currently the dominant methods in solving the problem of object maintenance in the video sequences are neural network approaches and discriminative correlation filters.

A characteristic feature of VOT-2018, unlike the previous ones, is that the best results in the categories (short-term, long-term, real-time, with short-term disappearances of objects) are shown by the algorithms built on the architectures of Siamese neural networks. The best compromise between reliability and processing speed can be achieved through the use of GPU.

Scientists use certain metrics to evaluate the performance of algorithms. In the simplest case, such a metric could be the proportion of samples by which the classifier made the right decision.

$$Accuracy = \frac{P}{N}$$

Where P is the number of correctly classified objects and N is the total number of objects.

In Table 1:

A – accuracy, reflects how accurately the algorithm determines the position of the object;

$$\varphi_t = \frac{A_t^G \cap A_t^T}{A_t^G \cup A_t^T}$$

where: A_t^T – define tracker predicted bounding box and A_t^G – define ground truth bounding box.

R – robustness, defines the amount of time when algorithm «loses» the object of interest;

EAO – expected average overlap, calculated as the average of the expected average overlap curve values over an interval of typical short-term sequence lengths.

²⁸ Kristan Matej (2018) The Sixth Visual Object Tracking VOT2018 Challenge Results.” *ECCV Workshops* (2018). The Sixth Visual Object Tracking VOT2018 Challenge Results. ECCV Workshops. URL : <http://votchallenge.net/publications.html>. DOI: 10.1007/978-3-030-11009-3_1. (in English)

Expected average overlap measure, computed as the average of the expected average overlap curve values over an interval $N_{lo} : N_{hi}$ of typical sequence lengths²⁹.

$$\epsilon = \frac{1}{N_{hi} - N_{lo}} \sum_{N_S=N_{lo}:N_{hi}} \epsilon \Phi_{N_S}$$

$$\Phi_{N_S} = \frac{1}{N_S} \sum_{i=1:N_S} \Phi_i$$

$$\epsilon \Phi_{N_S} = \langle \Phi_{N_S} \rangle$$

CPU is a central processing unit; GPU is a graphics processing unit;

Table 1

**Experimental results of the work
of the tracking algorithms on VOT-2018**

Algorithm	EAO	A	R	Controlled by CPU or GPU
LADCF	0.066	0.314	1.358	CPU
MFT	0.060	0.337	1.592	GPU
SiamRPN	0.383 1	0.586 1	0.276	GPU
UPDT	0.068	0.334	1.363	CPU
RCO	0.066	0.400	1.704	GPU
DRT	0.062	0.321	1.503	GPU
DeepSTRCF	0.063	0.418	1.817	GPU
CPT	0.081	0.479	1.358	GPU
SA Siam R	0.337 2	0.566 2	0.258	GPU
DLSTpp	0.125	0.514	0.824	GPU
LSART	0.055	0.386	1.971	GPU
SRCT	0.059	0.331	1.765	CPU
CFTR	0.062	0.319	1.601	GPU
CPT fast	0.152	0.515	0.726	GPU
DeepCSRDCF	0.062	0.399	1.644	GPU
SiamVGG	0.275	0.531	0.337	GPU
SA Siam P	0.286 3	0.533 3	0.342	GPU
CFCF	0.059	0.326	1.648	GPU
ECO	0.078	0.449	1.466	GPU
MCCT	0.061	0.359	1.742	CPU
CCOT	0.058	0.326	1.461	GPU

²⁹ Kristan M., Matas J., Leonardis A., Felsberg M., Cehovin L., Fernández G., Vojir T., Häger G., Nebehay G., Pflugfelder R.P. (2015) The Visual Object Tracking VOT2015 Challenge Results. *2015 IEEE International Conference on Computer Vision Workshop (ICCVW)*, pp. 564–586. DOI: 10.1109/ICCVW.2015.79. (in English)

Algorithm	EAO	A	R	Controlled by CPU or GPU
csrtpp	0.263	0.466	0.318	GPU
LWDNTthi	0.262	0.463	0.342	GPU
LWDNTm	0.261	0.455	0.323	S P G
R MCPF	0.064	0.329	1.391	GPU
FSAN	0.065	0.312	1.377	GPU
CSRDCF	0.099	0.477	1.054	CPU
DCFCF	0.080	0.321	0.665	CPU
UpdateNet	0.209	0.517	0.534	GPU
MBSiam	0.238	0.529	0.440	GPU
ALAL	0.067	0.404	1.667	GPU
CSTEM	0.239	0.472	0.379	CPU
BoVW CFT	0.063	0.331	1.615	CPU
C3DT	0.067	0.322	1.330	GPU
RSECF	0.074	0.414	1.569	GPU
DSiam	0.129	0.503	0.979	GPU
KFebT	0.195	0.475	0.670	CPU
MEEM	0.072	0.407	1.592	CPU
SiamFC	0.182	0.502	0.604	GPU
STST	0.156	0.466	0.763	GPU
DCFNet	0.180	0.471	0.548	GPU
DensSiam	0.174	0.462	0.688	GPU
SAPKLTF	0.117	0.481	0.946	CPU
Staple	0.170	0.530	0.688	CPU
ASMS	0.167	0.492	0.632	CPU
ANT	0.059	0.403	1.737	CPU
HMMTxD	0.073	0.416	1.564	CPU
DPT	0.126	0.483	0.899	CPU
STBACF	0.062	0.320	0.281 3	CPU
srdef deep	0.057	0.326	1.756	GPU
PBTS	0.102	0.411	1.100	CPU
DAT	0.139	0.436	0.749	CPU
LGT	0.059	0.349	1.714	CPU
RAnet	0.133	0.477	0.805	GPU
DFPReco	0.049	0.312	0.286	CPU
TRACA	0.136	0.424	0.857	GPU
KCF	0.134	0.445	0.782	CPU
FoT	0.130	0.393	1.030	CPU
srdef dif	0.061	0.398	1.925	GPU
SRDCF	0.058	0.377	1.999	CPU
MIL	0.069	0.376	1.775	CPU
BST	0.053	0.271	1.620	CPU
struck2011	0.093	0.419	1.367	CPU
BDF	0.093	0.367	1.180	CPU
Matflow	0.090	0.401	1.297	CPU
MRSNCC	0.060	0.328	2.088	CPU
DSST	0.077	0.396	1.480	CPU
IVT	0.065	0.386	1.854	CPU
CPOINT	0.057	0.290	1.901	CPU

Algorithm	EAO	A	R	Controlled by CPU or GPU
LIAPG	0.062	0.351	1.831	CPU
FragTrack	0.068	0.316	1.480	CPU
Matrioska	0.000	0.000	16.740	CPU

CONCLUSIONS

Given the above experimental results and considering the analysis of the results of VOT-2018, we came to conclusion that in the intellectual surveillance of the State border guard service of Ukraine to implement the predefined tasks, it is advisable to apply integrated approaches to the construction of algorithms of detection and tracking of objects in video sequences. Such complex approaches should make it possible to apply to the choice of the user the system algorithms built on neural network approaches or on the so-called traditional approaches (HOG, HAAR, Colornames, etc.).

Taking into account that tracking algorithms solve two problems: detection and tracking, and accordingly each of the stages must be studied and evaluated by different indicators, we consider it necessary to continue research in the following areas: analysis of algorithms for object detection and analysis of algorithms for tracking detected objects.

SUMMARY

One of the functions of intelligent video surveillance systems is the detection of objects crossing a conditional line on video sequences. Based on the analysis of research results of world leading scientists it was established that in order to for implement the above function it is necessary to solve the following problems: identifying objects, their localization, tracking from frame to frame and fixing the moment of crossing the group of pixels belonging to the «object of interest» with the pixels that belong to the «imaginary line».

The purpose of the article is to study the algorithms of object tracking in video sequences for further practical implementation in the processes of automatic processing of video information obtained from video surveillance cameras.

A review analysis of algorithms for object detection and tracking was performed, in the result of which the way to classify the algorithms of objects tracking was represented. Also, an experimental study of a number of tracking algorithms was conducted, as a result of which the recommendations were formed regarding the conditions of application of these algorithms for the purpose of their quantitative comparative evaluation and decomposition of detection and tracking problems.

As a result of the study, a classification of algorithms for object tracking was proposed. Recommendations on the conditions of application of these algorithms in real conditions were elaborated.

The analysis confirmed the importance of research algorithms for processing video information from video surveillance cameras and made it possible to identify the main directions of further research: analysis of algorithms for object detection and analysis of algorithms for tracking detected objects.

REFERENCES

1. Katerynychuk I., Babaryka A. (2018) Doslidzhennya algoritmiv viyavleniya ta suprovodzhennya ruhomih ob'ektiv u videoposlidovnostyah z kamer videosposterezhennya. [Analysis of the Technologies of the Functioning Specific Video Surveillance Systems and the Definition of their Directions of Improvement]. *Collection of scientific works of the National Academy of the State Border Guard Service of Ukraine. Series: Military and Technical Science*. no. 3(77). pp. 246–259 (in Ukrainian).
2. Zou Z., Shi Z., Guo Y., & Ye J. (2019) Object Detection in 20 Years: A Survey. ArXiv, abs/1905.05055. (in English)
3. Viola P. A., & Jones M. J. (2001) Rapid object detection using a boosted cascade of simple features. *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001, 1, I-I*. DOI: 10.1109/CVPR.2001.990517. (in English)
4. Viola P. A., Jones M. J. (2001) Robust real-time face detection. *Proceedings Eighth IEEE International Conference on Computer Vision. ICCV 2001, no. 2*, pp. 747–747. DOI: 10.1023/B:VISI.0000013087.49260.fb. (in English)
5. Dalal N., Triggs B. (2005) Histograms of oriented gradients for human detection. *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), 1*, 886-893 vol. 1. (in English)
6. Felzenszwalb P. F., McAllester D. A., Ramanan D. (2008) A discriminatively trained, multiscale, deformable part model. *2008 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1–8. DOI: 10.1109/CVPR.2008.4587597. (in English)
7. Redmon J., Divvala S.K., Girshick R.B., Farhadi A. (2015) You Only Look Once: Unified, Real-Time Object Detection. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 779–788. DOI: 10.1109/CVPR.2016.91. (in English)
8. Liu W., Anguelov D., Erhan D., Szegedy C., Reed S.E., Fu C., Berg A.C. (2016) SSD: Single Shot MultiBox Detector. *ECCV*. DOI: 10.1007/978-3-319-46448-0_2. (in English)

9. Lin T., Goyal P., Girshick R.B., He K., Dollár P. (2017) Focal Loss for Dense Object Detection. *2017 IEEE International Conference on Computer Vision (ICCV)*, pp. 2999–3007. DOI: 10.1109/ICCV.2017.324. (in English)
10. Zhou X., Wang D., Krähenbühl P. (2019) Objects as Points. *ArXiv, abs/1904.07850*. (in English)
11. Patil H., Bhagat K. (2015) Detection and tracking of moving object: A survey. *Int. Journal of Engineering Research and Applications*. vol.5 no. 11. pp.138–142. (in English)
12. Grandham Sindhuja, Renuka Devi. (2015) A survey on detection and tracking of objects in video sequence. *International Journal of Engineering Research and General Science*. vol. 3 no. 2. (in English)
13. Yilmaz A., Javed O. A., Shah M. (2006) Object tracking: A survey. *ACM Comput. Surv.*, vol. 38, no. 13. (in English)
14. Barga Deori and Dalton Meitei Thounaojam. (2014) A survey on moving object tracking in video. *International Journal on Information Theory (IJIT)*. vol. 3, no. 3 pp. 31–46. DOI: 10.5121/ijit.2014.3304. (in English)
15. Ann Maria Jacob and J Anitha (2012) Inspection of various object tracking techniques. *International Journal of Engineering and Innovative Technology*. vol. 2, no. 6, pp. 118–124. (in English)
16. Ojha S., Sakhare S. (2015) Image processing techniques for object tracking in video surveillance – A survey. *2015 International Conference on Pervasive Computing (ICPC)*, pp. 1-6. DOI: 10.1109/PERVASIVE.2015.7087180. (in English)
17. Grabner H., Grabner M., Bischof H. (2006) Real-Time Tracking via On-line Boosting. *BMVC*. (in English)
18. Lukezic A., Vojár T., Zajc L.C., Matas J.E., Kristan M. (2017) Discriminative Correlation Filter Tracker with Channel and Spatial Reliability. *International Journal of Computer Vision*, vol. 126, pp. 671–688. DOI: 10.1109/CVPR.2017.515. (in English)
19. Kalaz Z. (2011) Tracking-Learning-Detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, pp. 1409–1422. DOI: 10.1109/TPAMI.2011.239. (in English)
20. Kalal Z., Mikolajczyk K., Matas J.E. (2010) Forward-Backward Error: Automatic Detection of Tracking Failures. *2010 20th International Conference on Pattern Recognition*, pp. 2756–2759. (in English)
21. Henriques J., Caseiro R., Martins P., Batista J. (2012) Exploiting the Circulant Structure of Tracking-by-Detection with Kernels. *ECCV*. DOI: 10.1007/978-3-642-33765-9_50. (in English)
22. Henriques J.F., Caseiro R., Martins P., Batista J. (2015) High-Speed Tracking with Kernelized Correlation Filters. *IEEE Transactions on Pattern*

Analysis and Machine Intelligence, vol.37, pp. 583–596.
DOI: 10.1109/TPAMI.2014.2345390. (in English)

23. Bolme D. S., Beveridge J. R., Draper B. A., Lui Y. M. (2010) Visual object tracking using adaptive correlation filters. *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 2544–2550. DOI: 10.1109/CVPR.2010.5539960. (in English)

24. Babenko B., Yang M., Belongie S. (2009) Visual tracking with online Multiple Instance Learning. *2009 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 983–990. DOI: 10.1109/CVPR.2009.5206737. (in English)

25. Comaniciu D., Meer P. (2002) Mean Shift: A Robust Approach Toward Feature Space Analysis. *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, pp. 603–619. DOI: 10.1109/34.1000236. (in English)

26. Bradski G. R. (1998) Real time face and object tracking as a component of a perceptual user interface. *Proceedings Fourth IEEE Workshop on Applications of Computer Vision. WACV'98 (Cat. No.98EX201)*, pp. 214–219. DOI: 10.1109/ACV.1998.732882. (in English)

27. Held D., Thrun S., Savarese S. (2016) Learning to Track at 100 FPS with Deep Regression Networks. *ECCV*. DOI: 10.1007/978-3-319-46448-0_45. (in English)

28. Kristan Matej (2018) The Sixth Visual Object Tracking VOT2018 Challenge Results.” *ECCV Workshops (2018). The Sixth Visual Object Tracking VOT2018 Challenge Results. ECCV Workshops*. URL : <http://votchallenge.net/publications.html>. DOI: 10.1007/978-3-030-11009-3_1. (in English)

29. Kristan M., Matas J., Leonardis A., Felsberg M., Cehovin L., Fernández G., Vojir T., Häger G., Nebhay G., Pflugfelder R.P. (2015) The Visual Object Tracking VOT2015 Challenge Results. *2015 IEEE International Conference on Computer Vision Workshop (ICCVW)*, pp. 564–586. DOI: 10.1109/ICCVW.2015.79. (in English)

Information about the author:

Anatolii Babaryka,

Post-graduate student,

The National Academy of the State Border Guard Service of Ukraine named after Bohdan Khmelnytskyi, Khmelnytskyi, Ukraine,
46 Shevchenka str., Khmelnytskyi, 29007, Ukraine

ORCID ID: orcid.org/0000-0001-8534-7764