

# The matching points methods for solving the problem of the tracking objects

Polina Kovalenko, Andrey Mikhaylov, Alexander Kataev, Vladimir Rozaliev, Yulia Orlova

CAD Department

Volgograd State Technical University Volgograd, Russia

p.kovalenko.o@gmail.com, a.mikhaylov.v@gmail.com, alexander.kataev@gmail.com, vladimir.rozaliev@gmail.com, yulia.orlova@gmail.com

**Abstract** — This article interprets a concept of key points and its descriptors, object tracking and key points matching goals. The concept of object detection and detection classification methods are described. Key points matching methods based on standard OpenCV functions and matching machine learning methods are described. Comparative analysis of these methods working is represented by following parameters: standard deviation, maximum axis deviation, average operating time and execution accuracy.

**Keywords-** computer vision, detection, key point, descriptor, image, tracking.

## I. INTRODUCTION

Lately a lot of attention in computer vision area is paid to object tracking. It is a sequential definition of target object location for each stream frame. This technology is used in many areas related to video processing to obtain certain information.

Examples of systems using tracking technologies are sport events analysis systems, augmented reality systems, security systems. Also object tracking system can be installed on the quadcopter for object tracking from high. Police doesn't lose sight of culprit driving by car because of quadcopter flies after him.

The goal of this research is applicability definition of considered key point tracking methods for object tracking with quadcopter.

In spite of huge progress in computer vision researches, object tracking is still a difficult problem. It makes this research relevant.

## II. THE CONCEPT OF THE OBJECTS DETECTION

In matters of computer vision, the movement can be caused by the movement of the objects with a statically fixed camera, moving the camera relative to static objects, or by a combination of camera movement and tracking objects.

Tracking objects is the location of a moving object in time using the camera.

Detection is also called determination of availability of an object on an image and finding its position in the coordinate system of the pixels of the image.

The detection of moving objects is one of the components of the most tracking systems and real-time object movement analysis. Input data for such systems, as a rule, is a set of images representing a sequence of frames. The system should process and analyze this sequence of images.

## III. THE CLASSIFICATION OF THE METHODS OF DETECTION

The purpose of detection is to establish the presence of an object in the image. The detection methods allow you to analyze the scenes of the image and to get the information about the location of the objects on this scene.

Existing methods can be divided into the following categories: methods that use information for a characterizing individual points of an object (point features) to describe the object; object search methods that match a pattern-some description of the objects; methods of detecting the movement of objects - selection of the moving objects based on several images or video frames of the same scene.

## IV. OBJECT TRACKING

Object tracking is one of the most important problems of modern video analysis. In general, object tracking consists of object capturing, in other words to frame object in first stream frame image, and object frame transformation and shifting frame by frame [1].

Object can be described in different ways: bounding box or ellipse, a chain of united points etc. The Fig. 1 shows bordered frame around target object [2].

The main problem of such a task is the determination of the object's changes due to the presence of occlusions, noise on the image, changing the shape, color and illumination of the object itself during the movement. Occlusions occur when some objects overlap, or when the target disappears from the frame. This can lead to incorrect operation of tracking algorithms. Noise is a phenomenon in which an object is hardly distinguishable from the background. Examples of such changes are shown in Fig. 2.

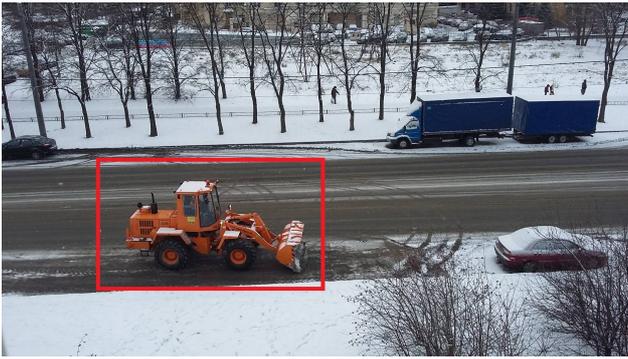


Fig. 1. Tracking object is marked with frame

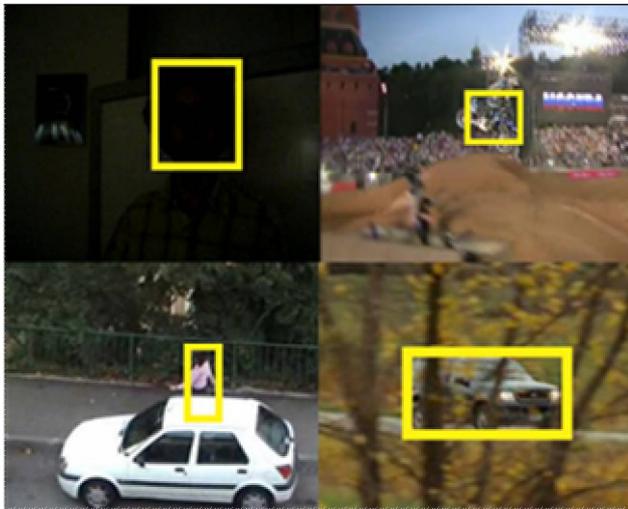


Fig. 2. The examples of strong object changes are: a light variation, a difficult separation from the background, occlusions

### V. KEY POINTS AND ITS DESCRIPTORS

Particular importance in the image description construction of an image has the image specific parts detection, for example corners, edges, intensity extremum regions [3].

Scene key point or point feature is an image point as a pixel with a characteristic neighborhood i.e. neighborhood is different from other points neighborhood. Calculated attributes vector based on intensity, gradient or another characteristics describes the key point.

Using a key point, it is possible to analyze both whole images, and objects on them. A good key points help to change scale, view and overlap of scene or object [4].

Key point descriptor is a numerical vector which characterizes image feature in the point neighborhood [3].

### VI. KEY POINTS MATCHING

Image key points detection and matching is an important computer vision problem.

At the first sight, a set of appropriate points on the images provides not enough in-formation about image and observable scene but it is not like that.

For example, if you have some images from one scene and a set of appropriate points on these images you can define camera settings and location for each image [5].

Consider the problem in more detail. There are two scene images. It is necessary to find a set of point pairs from both images in the way that points correspond the same object point on images (Fig. 3).

However not all points are suitable for matching. It is difficult to find a point corresponded to homogeneous surface point. Therefore, key points are used for matching.

This process consists of three stages: key points selection through feature detector; key points describing through feature description; receiving a set of key point conformities through matcher.

This process is not so complicated in virtue of the existence of large number of computer vision libraries. OpenCV library is most prominent, it is implemented on C++ and has more functions than other ones. There is Feature Points Detection & Description module in OpenCV, which contains implementations for different detectors, descriptors and matchers.

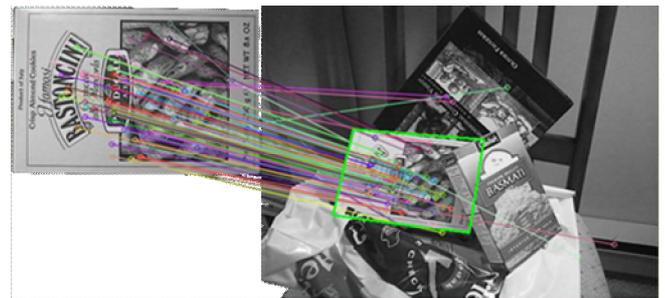


Fig. 3. Comparison of keypoints on the two images.

### VII. KEYPOINTS MATCHING BASED ON STANDARD OPENCV FUNCTIONS

#### A. Key points matching without filtering

This method uses OpenCV matcher for finding accordance between two sets of image points [6]. Matcher receive as input the image of search object and image to find this object.

As a result, accordance between object points and non-object points will be constructed.

At first key points searching and descriptors computing are processed for both images. Images provide found key points description through estimating its location with description of surrounding areas. Then points matching is executed.

Known algorithms for finding matches between two images key points often give many outliers.

It leads to the fact that algorithms cannot match test images correctly, even images coincide and are made from different angle (Fig. 4).

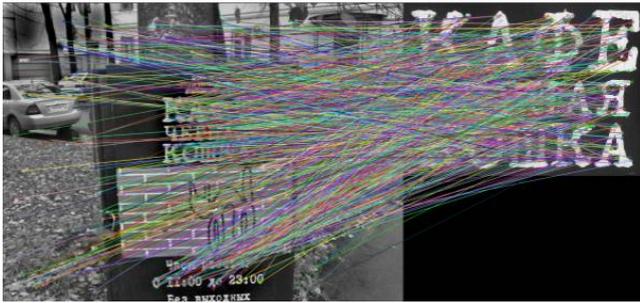


Fig. 4. Example of incorrect comparison the two images.

*B. Key points matching with filtering*

This method is a modification of the previous one. Sifting coincided points, which refer to the same key point, improves an accuracy. It performs in the following way: minimum and maximum distance between points are computed after matching. Distance between selected point pairs must be less than half-sum of maximum and minimum distances between them.

This method shows better results than method above.

**VIII. KEY POINTS MATCHING WITH MACHINE LEARNING**

One possible approach of solving the object tracking through key points problem is a usage machine learning algorithms for building an object classes model and output algorithms for object searching on the image.

Model building consists of two stages [7]: feature extraction specific for class object, characteristic feature vector building for object key points (corners, edges or contour of object) or for whole object; model training on the obtained features for following object recognition.

*A. Matching through Support Vector Machine*

This method based on machine learning used for classifier creation. Classifier recognizes what features belong to the object.

Support Vector Machine (SVM) belongs to boundary methods group. Methods of this group define classes through areas boundaries [8].

After key points searching and its descriptors computing for the same image it is necessary to train SVM with training sample. Key points descriptors represent the training sample.

The first class – all points belong to the object.

The second class – all points don't belong to the object.

SVM classifier is trained on test sample after its preparation.

Then key points searching and its descriptors computing are performed for each frame.

OpenCV library function predict() is invoked for each current frame descriptor. This function defines appropriate point belonging to any class.

*B. Matching through neural network*

This method is similar to the previous one. It also based on machine learning namely neural networks.

Neural networks are widely used for different problems solving classification is one of them.

Neural network training process is configuring its internal parameters for specific task. Training and prediction processes are the same as in SVM.

**IX. THE RESEARCH METHODOLOGY DESCRIPTION**

The research was conducted in Microsoft Visual Studio 2012 Ultimate using OpenCV 3.0 library.

Described methods comparative analysis was carried out for checking key points methods are suitable for object tracking through the quadcopter. Comparison criteria are standard deviation the object center coordinates from real object center coordinates (in pixels), average algorithm operating time for single frame (in seconds), maximum the object center coordinates deviation from real object center coordinates (in pixels), algorithm accuracy (in percent).

SIFT and SERF key points searching and descriptors computing algorithms are used in research.

As a test data MPEG-4 video was recorded and text file with object location information for each frame was prepared. This input data set allows to calculate tracking methods execution statistics.

The yellow balloon with two eyes and smile is a tracking object for test video (Fig. 5). Fig. 6 represents a test video frame with visualized SVM algorithm performing. Points belonged to the tracking object are blue other points are green.



Fig. 5. Tracked object

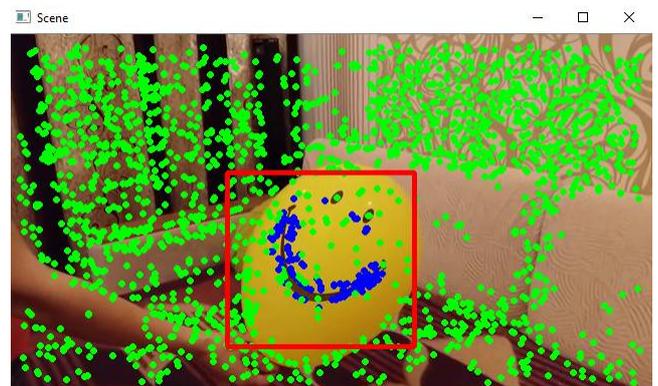


Fig. 6. Frame from a test video.

## X. RESEARCH RESULTS

Table 1 contains key points tracking methods execution results.

TABLE I. KEY POINTS OBJECT TRACKING METHODS EXECUTION STATISTICS.

Method of object tracking	The descriptor keypoints	The standard deviation (in pixels)	Max. deviation of the X, Y (in pixels)	The average running time per frame (in seconds)	The accuracy of the algorithm (in percent)
Matching without filtering	SIFT	7.58	X: 17.13 Y: 120.78	1.31	72
	SURF	31.27	X: 51.52 Y: 384.12	0.43	66
Matching with filtering	SIFT	9.23	X: 14.21 Y: 23.13	1.32	76
	SURF	10.11	X: 7.32 Y: 98.22	0.24	69
Matching through SVM	SIFT	17.21	X: 56.43 Y: 85.11	0.48	79
	SURF	19.11	X: 46.19 Y: 154.44	0.29	87
Matching through ANN	SIFT	45.62	X: 59.53 Y: 57.71	8.45	38
	SURF	31.47	X: 55.40 Y: 109.63	2.38	52

SVM algorithm has the faster execution and is the most accurate method. Average execution time for single frame is 0.29 seconds and accuracy is 87%. The matching through neural network (ANN) has the slowest execution – 8.45 seconds and its accuracy is the lowest – 38%.

Additional advantage of SVM is a case of continuity object tracking even after losing tracing object. Machine learning algorithms facilitate this. In addition, SURF key point searching and descriptors computing algorithms are executed much faster than SIFT.

The key points cut-off threshold decreasing increases the operating time of algorithms based on machine learning but at the same time accuracy is increased too.

## XI. CONCLUSION

As a result of this work different key point matching methods were investigated: based on standard OpenCV functions method, neural network and support vector machine method.

Methods executed with real test data and received information was used for identifying and analysis statistics.

Thus, the conclusion is that methods based on key points are not suitable for object tracking through quadcopter because its performance is too low.

## ACKNOWLEDGMENT

This work was partially supported by RFBR (grants №15-37-70014, 16-07-00407, 16-47-340320).

The research work carried out as part of the grant to the winner of the program «UMNIK MIPT-autumn 2015» under the contract № 9886/2015 of 10.03.2016 (code 0020827) with the project «Development of methods for visual search and track targets using quadcopters».

## REFERENCES

- [1] A. A. Shaposhnikov, E. V. Shaposhnikova and A. I. Shaposhnikov, "K voprosu o kachestve raboty algoritmov slezheniya za obyektami," *Komp'yuternyye issledovaniya i modelirovaniye*, 2014, T. 6, vol. 4, pp. 495-502.
- [2] D. A. Vrazhnov, A. V. Shapovalov, V. V. Nikolayev, "O kachestve raboty algoritmov slezheniya za obyektami na video," *Komp'yuternyye issledovaniya i modelirovaniye*, 2012, T. 4, vol. 2, pp. 303-313.
- [3] P. N. Polovinkin, *Detektory i deskriptory klyuchevykh toчек. Algoritmy klassifikatsii izobrazheniy. Zadacha detektirovaniya obyektov na izobrazheniyakh i metody yeyo resheniya*, NNGU, Nizhniy Novgorod : Izd-vo NNGU, 2013, 38 p.
- [4] Y. Amit, "2D Object Detection and Recognition: models, algorithms and networks," *The MIT Press*, 2002, 325 p..
- [5] V. Gaganov, "Invariantnyye algoritmy sopostavleniya tochenykh osobennostey na izobrazheniyakh," *Setevoy Zhurnal Grafika i Mul'timedia*, 2009, vol. №7(1), pp. 57-69.
- [6] V. D. Kustikova, "Razrabotka mul'timediynykh prilozheniy s ispol'zovaniyem bibliotek OpenCV i IPP," NNGU, Nizhniy Novgorod : Izd-vo NNGU, 2013, 34 p.
- [7] B. Apolloni, *Machine Learning and Robot Perception*, Springer, 2005, 900 p.
- [8] A. D. Lun'kov and A. V. Kharlamov, "Intel'lektual'nyy analiz dannykh," SGU, Saratov : Izd-vo Saratovskogo Universiteta, 2012, 96 p.