International Conference Engineering Technologies and Computer Science: Innovation and Application (EnT-2020), June 24-27 (October 6-8), Moscow - Saint-Petersburg, Russia

Adaptive Algorithm For Object Tracking In Video Image

Andrey V. Shipatov, Evgeni M. Portnov, Aung Kyaw Myo, Naing Linn Aung

National Research University of Electronic Technology (MIET)
Institute of Systems, Software Engineering and Information Technology

SUMMARY

- Introduction
- Objective of the work
- Tracking algorithm
- Cascade detector
- Experimental results
- Conclusion

INTRODUCTION

The use of algorithms for automatic tracking of objects on video images as a part of on-board control systems can significantly increase the autonomy and efficiency of various types of unmanned vehicles.

The main requirement for the tracking algorithm is the reliability and stability of tracking under conditions of optical interference, partial or complete overlap of the object with the background, sharp movements of the camera, the object goes out of the viewport and then appears in the tracking window after a certain time interval.

The reason for the failure of tracking is often a strong change in the image of the object, which occurs, for example, when the camera angle of the unmanned aerial vehicle (UAV) is sharply changed or is characteristic of video images obtained from thermal imaging cameras. Therefore, the use of preliminary information about the properties of the object (its shape, texture, distribution of pixel intensities) is not possible. Also, one of the key requirements for the tracking algorithm that works as part of the UAV control system is to work in real-time, without the ability to analyze the entire video sequence before the start of tracking. Thus, the task of developing an algorithm for sustainable tracking of objects arises in real-time with minimal a priori information.

OBJECTIVE

This paper presents a method for solving the problem of stable tracking of an object with a changing image, based on a universal approach to the construction of adaptive tracking algorithms, which includes three components: tracking, training and detection.

- The object's trajectory is tracked by a certain short-term tracking algorithm, which works parallel with the detector, allowing it to be re-initialized after a failure.
- The results of the tracking algorithm are used to build a model of the object and partially controlled training of the detector using two independent processes that continuously track the detector errors and update it to avoid these errors in the future. Since the detector is a binary classifier that relates the given image fragment to the object ("positive" fragment) or the background ("negative" fragment), it generates two types of errors: false-negative fragments and false-positive fragments.
- Each of the two training processes identifies classification errors of the corresponding type and corrects the detector by introducing its errors, which are mutually compensated so that the learning process is sustainable. Such a method of dynamic classifier training is called P-N-training.

TRACKING ALGORITHM

The problem of tracking an object is formulated as a problem of Bayesian estimation of the area occupied by it, which is reduced to the equation of evolution of a contour describing the boundary of the object and represented by a zero-level line of a two-dimensional function called the characteristic function.

Let be I^k, $k = \{0, 1, 2, \ldots\}$- a sequence of images occupying the region $\Omega \subset R^2$, and $R_0 \subset \Omega$, $R_1 \subset \Omega$ - the areas occupied by the object in the image I^n and I^{n+1} at the moments n and $n+1$ respectively. The estimation of the unknown area R_1, occupied by the object in the frame I^{n+1} is based on the known ones I^n, I^{n+1} and R_0 is performed by the method of a posteriori maximum:

$$\hat{R}_1 = \arg \max_{\mathbf{R} \subset \Omega} P\left(R_1 = R \middle| I^n, I^{n+1}, R_0\right) \qquad (1)$$

By using the Bayes rule, (1) can be rewritten in the form:

$$\hat{R}_1 = \arg \max_{R \subset \Omega} P\left(I^{n+1} \middle| I^n, R_0, R_1 = R\right) P\left(R_1 = R \middle| I^n, R_0\right) \qquad (2)$$

CASCADE DETECTOR

To detect an object in the image after the tracking algorithm fails, a detector was used in this work, which is a cascade of classifiers: a combination of classifiers, in which the output of one classifier is input to another. The input image is scanned by a sliding window, and for each position of the window, a decision is made about the presence or absence of an object in the corresponding fragment. At the first stage, more than 50% of the background fragments are discarded by the threshold pixel dispersion filter of the fragment. The remaining fragments are classified by using an ensemble of basic classifiers based on the random forest. In each fragment pre-processed with a Gaussian filter to reduce the impact of noise and random movement, all the base classifiers perform a series of pairwise comparisons of the intensities of randomly selected pixels. The result of each comparison is the value 0 or 1, and the binary code x corresponds to the set of comparisons of the i^{th} tree, which defines the index in the array of posterior probabilities, $P_i(y|x)$, where $y \in \{0,1\}$ - is the label of the fragment class (0 - background, 1 - object). The posterior probability of the i^{th} base classifier is estimated as follows:

$$P_i(y|x) = \frac{N_p^i(x)}{N_p^i(x) + N_n^i(x)}$$

where: $N_p^i(x)$ - number of positive fragments,

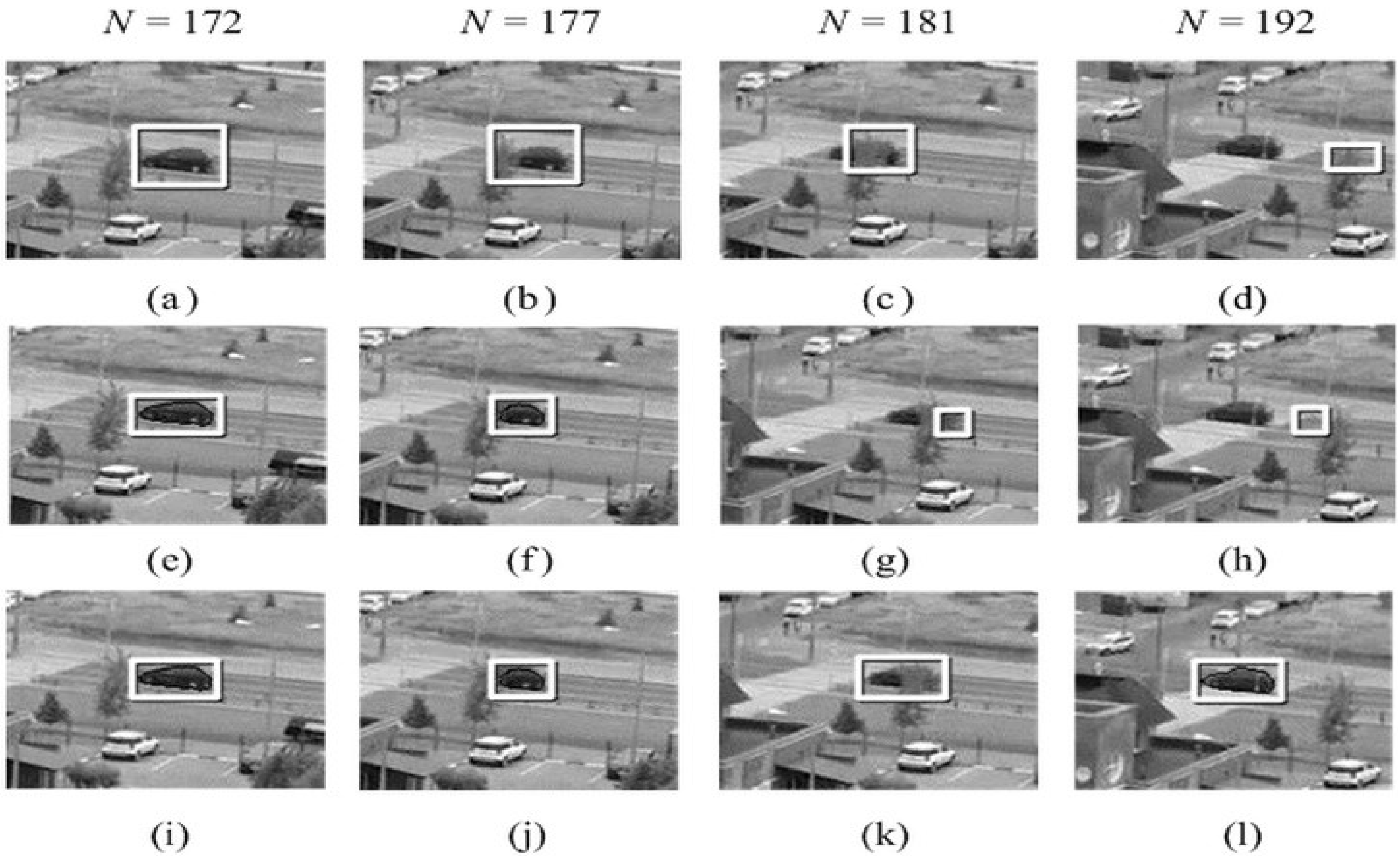
$N_n^i(x)$ - number of negative fragments, to which the code x had corresponded.

Then the average posterior probability is calculated for all the ensemble classifiers, and if it exceeds 50%, it is assumed that the fragment contains an object, otherwise, it belongs to the background.

The interaction of the components of tracking algorithm

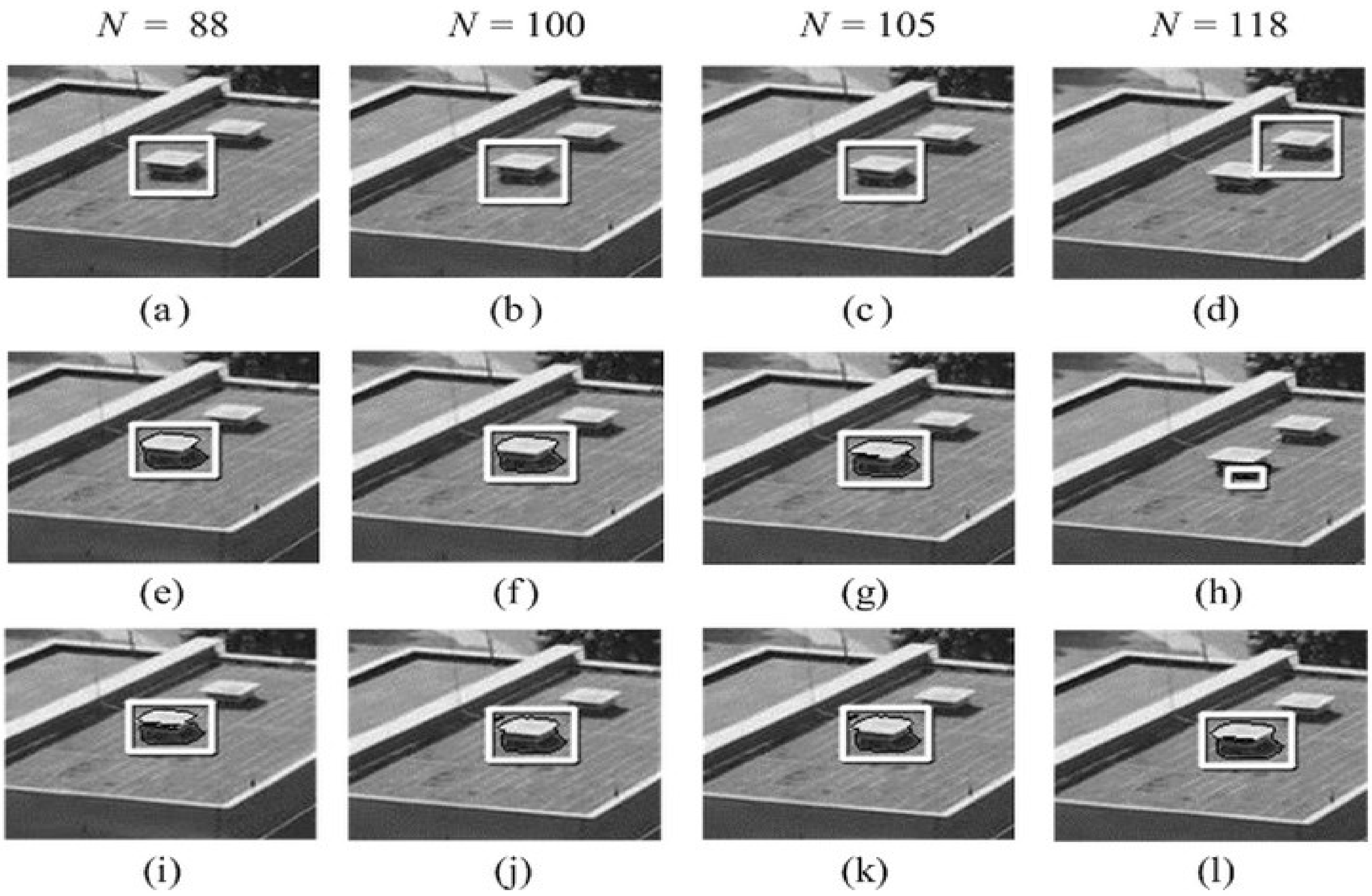
- The characteristic function $u(x,t)$ of the tracking algorithm is initialized by the sign distance function [5], constructed for the object outline selected at the initial stage. In each frame, the adaptive tracking algorithm estimates the area occupied by the object and presents the prediction result in the form of a minimal rectangle containing the object.
- This prediction is supplemented by the hypotheses of a detector processing each frame together with a tracking algorithm, and as a result, a rectangular frame is determined that most reliably describes the position of the object.
- The resulting frame is used to generate positive and negative training examples, where P-N training of the detector are occurred [2], [3]. In the event of a short-term tracking algorithm failure, the position of the object is estimated only based on the most reliable hypothesis of the detector.
- Restoration of the object contour after tracking failure is performed using the geometric method of active contours described in [6] and used in the developed automatic tracking system to highlight the object when the operator hovers over it.
- The rectangular frame of the region detected by the detector sets the initial approximation of the object contour and is used to initialize the characteristic function $u_0(x)$ of the object contour extraction algorithm. The selected contour allows reinitializing the tracking algorithm by defining its characteristic function using the sign function of the distance to the contour.
- If the tracking algorithm fails and the detector also does not detect any positive fragments, it is assumed that the object in this frame is missing.

The influence of interference (overlapping by a background object) on the vehicle tracking on a television video



TLD algorithm with a tracking component based on the LK (a - d) method; proposed adaptive short-term tracking algorithm (e - h); TLD algorithm with the proposed tracking algorithm (i - l), N - frame number

Test results of tracking algorithms on a video image obtained from a quadrocopter, in the presence of similar objects



TLD + LK (a - d) algorithm; proposed adaptive short-term tracking algorithm (e - h);
TLD algorithm with the proposed tracking algorithm (i - l), N - frame number

Conclusion

- A similar result was obtained for a video image obtained from a camera mounted on a quadrocopter suspension The presence of two similar images confuses the TLD + LK algorithm, resulting in the interception of another object.
- The short-term tracking algorithm based on Bayesian estimation of the area occupied by the object loses it without the possibility of recovery.
- The proposed algorithm through to the detector can restore tracking of the object after the failure of the tracking algorithm and demonstrates great stability in the presence of similar objects.
- This algorithm was used by the authors in the software package of the developed system of automatic tracking of objects in video images, which is part of a universal open software and hardware platform for designing video streaming processing devices for UAV environmental monitoring.

Acknowledgment

This work was financially supported by the Ministry of Science and Higher Education of the Russian Federation (Unique identifier of the project RFMEFI60519X0181).

References

- Lucas B.D., Kanade T. An iterative image registration technique with an application to stereo vision // International Joint Conference on Artificial Intelligence. 1981. Vol. 81. P. 674–679.
- Kalal Z., Matas J., Mikolajczyk K. Tracking-Learning-Detection // IEEE Transactions on Pattern Analysis and Machine Intelligence. January 2010. Vol. 6, № 1.
- Kalal Z., Matas J., Mikolajczyk K. P-N learning: Bootstrapping binary classifiers by structural constraints // CVPR 2010. P. 49–56.
- Mansouri A.R. Region tracking via level set PDEs without motion computation // IEEE Transactions on Pattern Analysis and Machine Intelligence. 2002. Vol. 24, № 7. P. 947–961.
- Osher S., Fedkiw R. Level Set Methods and Dynamic Implicit Sur-faces // New York : Springer-Verlag, 2002. 273 p.
- Li C., Xu C., Gui C., Fox M.D. Level Set Evolution Without Re-initialization: A New Variational Formulation // CVPR 2005. Vol. 1. P. 430–436.

Thanks For Your Attention!