FAST METHOD OF MOVING OBJECT TRACKING

Kamil Ekštein

Laboratory of Intelligent Communication Systems
Department of Computer Science and Engineering
Faculty of Applied Sciences
University of West Bohemia
Plzeň, CZECH REPUBLIC
E-mail: `kekstein@kiv.zcu.cz`

Abstract: The article describes a series of tests of different computer vision techniques used to track a moving object in a scene taken by a digital camera (e.g. DV camcorder). The performance provided by standard techniques was found unsuitable either because of high computational demands (or complexity respectively) with respect to expected target real-time application, or for poor movement tracking ability. A novel approach was developed with focus on high performance in moving object tracking and low computational demands so that the method can be easily implemented into low power systems (embedded computers based on 80386 processors, high-level microcontrollers, etc.).

Keywords: computer vision, motion detection, object tracking

1. INTRODUCTION

The research was motivated by a need of a set of computer vision techniques for an autonomous mobile system navigation. The basic idea behind the navigation and autonomous control of a mobile system is that the moving objects are considered important as they could cause a collision in a short-time horizon. On the other hand the still objects (merged to the scene background) present no collision danger within short time periods. The mobile system is navigating through the space by an analysis of the scene – the global trajectory planning is usually not time critical. However, considering other moving objects in the space, the autonomous system must be able to detect a collision threat and plan a dodging manoeuvre in the shortest time possible.

Another field where it is well possible to benefit from a fast and efficient movement detection and analysis method is automatic surveillance used in various security applications.

2. TESTED STANDARD METHODS AND THEIR PERFORMANCE

During an extensive testing phase, 3 computer vision methods (nowadays considered standard according to (V. Hlaváč, 1992)) were tested within the framework of the moving object tracking task. All of them were used to perform an image segmentation[1]: (i) *Hough transform*, (ii) *SUSAN feature detector*, and (iii) *Skeletonisation*.

[1] Subsequently, positions of the detected objects are analysed and compared on consecutive frames. Such a motion analysis can be considered independent on the used segmentation method.

Hough transform proved to have expectedly outstanding ability to find an analytically described object (in fact a geometrical shape). Its sensitivity to image noise is very low and its ability to cope with incomplete data is also high above average. These desirable features are balanced by high computational demands of the method. Although the testing was performed on a powerful dual-core Pentium machine, and the algorithm was written in C++ using speed optimisation, it was not suitable for a real-time application, as the resulting frame rate was 3 fps.

The **SUSAN (Smallest Univalue Segment Assimilating Nucleus) feature detector** is quite a novel approach presented some ten years ago by Smith and Brady in (S. M. Smith, 1995). It has a very low computational load and is easy to implement (so that it is suitable for small, low-consumption embedded systems, microcontrollers, and other poor performance hardware). On the other hand this method was not able to overcome substantial amount of image noise and thus produced a lot of false features that had no connection with the actually tracked moving object.

Skeletonisation proved good performance when working on noisy images and incomplete data. The computational load was significantly lower compared to the Hough transform. However, the segmentation ability dropped congruently – especially small still objects of the same or similar geometrical shapes caused confusion. Moreover such a method needs further processing of the resulting data to segment the scene so that the moving object can be localised.

3. THE METHOD: SEGMENTATION VIA TEMPORAL DIFFERENCE

As the previously mentioned methods did not fully satisfy the conditions, an alternative approach was developed and tested: The image is not segmented in order to locate objects but to locate a movement. The scene images are transformed using the following FIR spatiotemporal filter:

$$y(i,j,t) = \psi(i,j,a) \cdot |x(i,j,t) - x(i,j,t-\varphi(a))|,$$

where i and j are pixel coordinates within a single image, t is time coordinate within an image stream, i.e. t-th image is processed, x is an input image function, y is an output image function, ψ is a non-linear smoothing function, and φ is a temporal shift function. Both ψ and φ are dependent on a smoothing parameter a.

The ψ function is also dependent on spatial coordinates i and j that enables to prefer a region where the tracked object was localised in the past (presuming the object is moving at certain speed that can't change arbitrarily fast due to object inertia).

The form of the filter formula suggests that the filter is a form of a numeric *first-order temporal derivative*. The basic property of a derivative, i.e. the ability to express and measure the change, is used within the proposed method to detect the movement. Figure 1 shows the difference image, in fact a visualised output of the derivative. The filter produces an image where all the pixel intensities are given by a difference between the corresponding pixels in the consecutive images. The pixels that represent only a small change (expressed by the smoothing parameter) are merged into the background. The pixels that have the intensity above the threshold form objects that are called *movement traces* as they represent those areas of the image where a movement appeared.

After the filtering, standard thresholding and labelling procedures are applied to localise, count and analyse the movement traces. These are analysed using vector field analysis techniques.

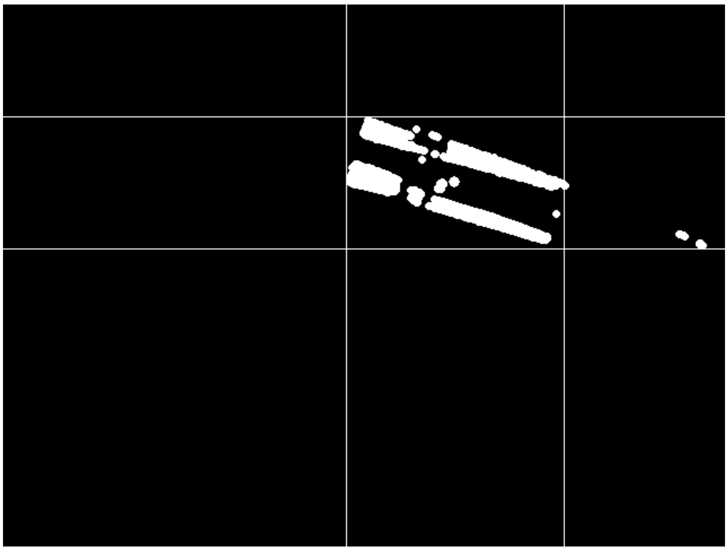

Fig. 1: Visualised temporal difference between two consecutive images

4. APPLICATION

The developed method can be used in various technical applications: (i) navigation of an autonomous mobile system (e.g. AI driven car, aircraft, etc.) where it can act as a part of a collision-avoiding subsystem; (ii) security appliances like e.g. aiming a security camera around the observed area so that an intruder stays locked in the centre of take so that his/her mischief activity can be monitored; (iii) military appliances like e.g. locking aircraft finders; (iv) robotics where it can be used as a part of a robot navigation system – nowadays e.g. robotic football is gaining popularity and in such a task it is important to detect fast movements of a ball and opponent robots; etc. Many other technical applications can be found.

The following figures 2 and 3 show the window of the testing application O'Trac. Two different types of scenes with moving objects are processed. Figure 2 shows an aerial scene where there is an aircraft in the field of vision of the system. Such a scene is typical for e.g. locking finders connected to weapon-aiming systems, etc. Figure 3 shows a typical scene of a security camera. The goal in this kind of applications is to alarm an armed force that can capture the intruder or at least to keep the focus on him or her.

5. RESULTS

A simple and effective method of moving object tracking was developed. The method is easy to implement, has very low computational power demands and is thus very suitable for low power computer systems (e.g. 80386- or ARM-based embedded computers). The method is reliable and stable in tasks where the tracked object moves at speeds sufficiently higher than the speed of change of the monitored scene. The most substantial disadvantage of the presented method is a complete inability to indicate an unexpected disappearance of the tracked object. Detailed analysis can be found in (Weinfurt, 2006).

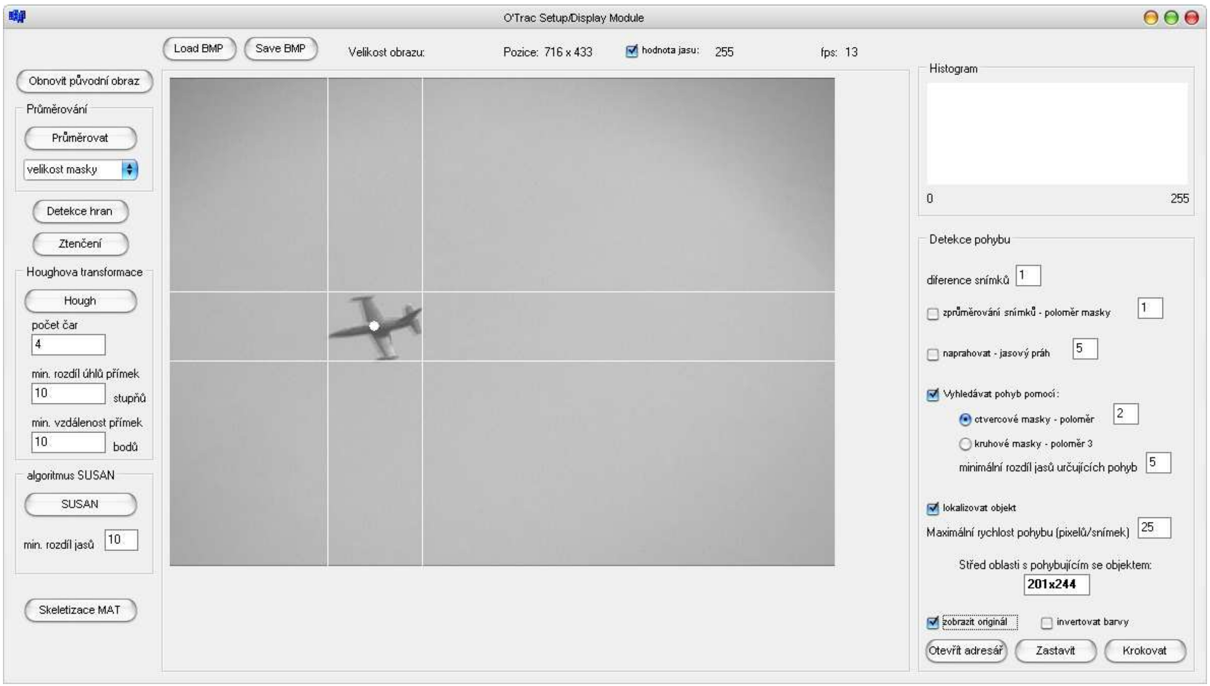

Fig. 2: Testing application window — locating moving object in a typical aerial scene

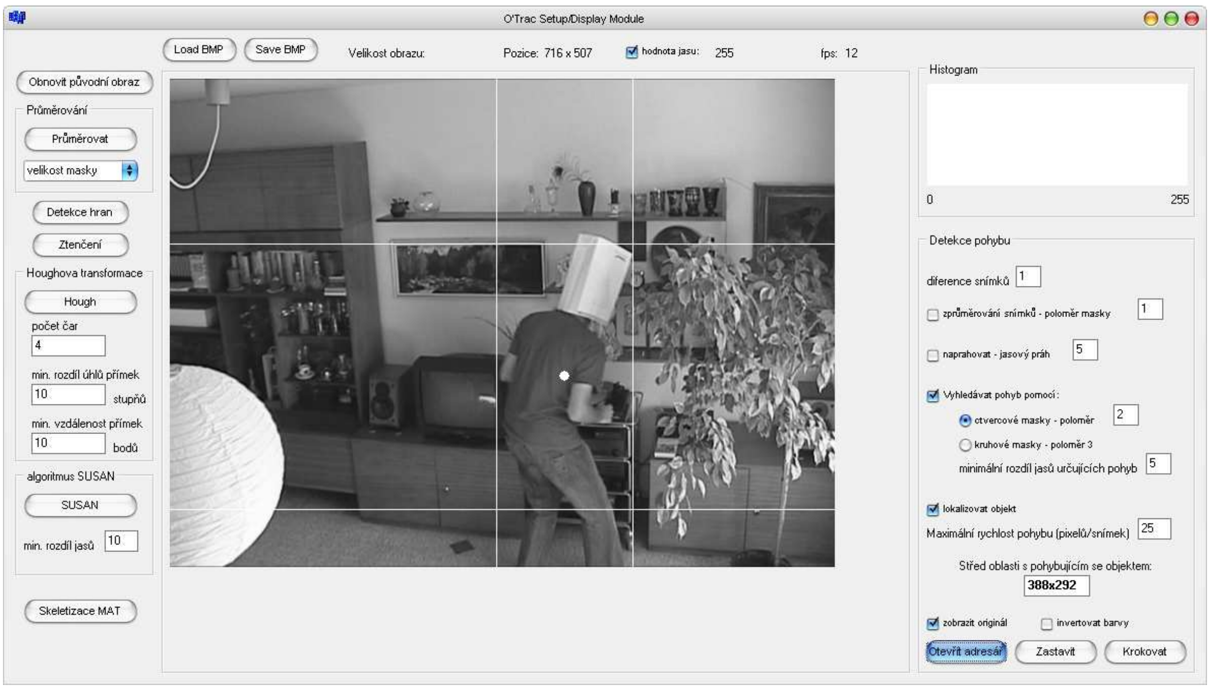

Fig. 3: Testing application window — locating moving object in a typical security camera scene

ACKNOWLEDGEMENTS

This work was supported by an internal grant of the Department of Computer Science and Engineering.

REFERENCES

S. M. Smith, J. M. Brady (1995). Susan — a new approach to low level image processing. Technical Report TR95SMS1c. Oxford Centre for Functional Magnetic Resonance Imaging of the Brain, and Department of Engineering Science, Oxford University, Oxford, UK.

V. Hlaváč, M. Šonka (1992). *Počítačové vidění*. Grada. Prague, Czech Republic.

Weinfurt, P. (2006). Detekce a analýza pohybu objektu ve scéně (detection and analysis of object movement in a scene). Master's thesis. Dept. of Computer Science and Engineering, Faculty of Applied Sciences, University of West Bohemia. Plzeň, Czech Republic.