

AN EFFICIENT SENSOR FOR TRAFFIC MONITORING AND TRACKING APPLICATIONS

Based on Fast Motion Detection at the Areas of Interest

Nikolaos Zournis-Karouzos

Department of Electrical and Computer Eng., Aristotle Univ. of Thessaloniki, Thessaloniki, Greece

Alexandra Koutsia, Kosmas Dimitropoulos, Nikos Grammalidis

Informatics and Telematics Institute, CERTH, 1st km Thermi-Panorama Rd, Thessaloniki, Greece

Keywords: Motion detection, traffic monitoring, target tracking, background extraction and update, A-SMGCS.

Abstract: We propose a novel video sensor for real-time motion detection at specific user-defined regions of interest, designed primarily for traffic monitoring, surveillance and tracking applications. Specifically, the new sensor a) supports virtual detectors with a generalized (polygonal) shape, thus providing additional flexibility in the design of detector configurations, b) is based on fast implementations of recent state-of-the-art background extraction and update techniques and c) constitutes a generic, inexpensive software solution, which can be used with any video camera. First experimental results confirm that the new video sensor meets the expectations in terms of real-time performance and demonstrates the additional functionalities, according to which it was designed. The final goal is to use this new sensor as an alternative, improved version of embedded motion detection video sensors (like Autoscope®).

1 INTRODUCTION

Within the last years, there is increased market demand for the development of efficient automated systems that use computer vision techniques for real-time traffic monitoring, surveillance and accident control. These systems have also been used to augment existing Advanced Surface Monitoring, Guidance and Control Systems (A-SMGCS), (ICAO document, 1986) at airports (Besada et al, 2005), (Pavlidou et al, 2005).

An example of such a system is Autoscope® Solo Wide Area Video Vehicle Detection System. However, such systems are usually very expensive, since they use specialized cameras with additional integrated (onboard or not) hardware/software for real-time motion detection. Furthermore, they are not very efficient if used as sensors for tracking or security surveillance applications. In the FP5 IST INTERVUSE and FP6 EMMA projects (Pavlidou et al, 2005), Autoscope sensors were successfully used to provide an alternative A-SMGCS solution for small-medium airports without any A-SMGCS

means or to augment an existing A-SMGCS system (typically based on a surface radar) by covering specific “blind spots” (usually occurring near buildings or other obstacles).

However, specific shortcomings were identified: a) constraints due to the rectangular nature of virtual detectors, b) use of older, traditional image processing algorithms and c) the high cost of video sensors. This paper proposes a novel system for real-time motion detection at specific regions of interest within the camera’s field of view, which aims to avoid the above shortcomings. Specifically, it a) is using virtual detectors of a generalized polygonal shape, b) is based on fast implementations of recent state-of-the-art background extraction and update techniques and c) is inexpensive, being implemented entirely in software. First experimental results confirm that the new video sensor meets the expectations in terms of real-time performance. The final future goal is to use this new sensor as an alternative, improved version of the Autoscope video sensors for the targeted applications.

The rest of this paper is organized as follows: In Section 2, a brief introduction is made to the

Autoscope Vehicle Detection and its use withing the INTERVUSE project. Section 3 presents the four background extraction and update techniques extended in this paper to provide fast and reliable motion detection for the generalized-shape (polygonal) virtual detectors. Finally, Section 4 contains experimental results and conclusions demonstrating the computational gains achieved by the proposed technique.

2 THE AUTOSCOPE VEHICLE DETECTION SYSTEM

The Autoscope® Solo Wide Area Video Vehicle Detection System is an advanced, sophisticated, traffic surveillance system that uses machine vision technology to produce highly accurate traffic measurements (Image Sensing System, 2007), (Michalopoulos et al, 1993). It is used for traffic control centres and Internet information systems as well as incident detection to improve emergency response times of local authorities.

The Autoscope camera has a built-in Machine Vision Processor (MVP) which provides many benefits such as: a) there is no need for high bandwidth video transmission between the camera and the MVP, b) enables closed loop control of the camera optics such as illumination, gain, brightness and electronic zoom by the vision processor itself, c) makes the system more easily portable.

Autoscope cameras are addressable by a unique IP address and can be linked to each other, as well as to a PC for configuration and statistics collection, using RS-485 communication.

Each camera can detect traffic in multiple locations within its field of view. Rectangular areas, called virtual detectors (VDs), can be defined by the user on the camera image plane, each corresponding to a binary output. More complex virtual detectors can also be defined by combining detector outputs by means of logical and mathematical expressions (AND, OR, NOT, time based consideration, averages, sums, etc). The main advantage of the use of virtual detectors is that processing involves only the pixels of the specified areas of the image, thus reducing the computational requirements.

Once the location of virtual detectors has been specified, the background in the absence of vehicle is estimated. Virtual detectors detect the presence of vehicles by estimating the statistics of the background from which a threshold is determined. Then the instantaneous image pixel values are

compared with this threshold and if they are greater it means that a vehicle is present (Michalopoulos, 1991). Over time, the inbuilt pattern recognition software learns pattern of contrast, thus very well coping with night, fog, snow and rain, as experience with road traffic has shown.

Within the INTERVUSE project (Pavlidou et al, 2005) this detection system was used for monitoring of airport ground traffic. More specifically, information from all available virtual detectors configured in the video sensor network is continuously provided to Video Sensor Data Fusion (VSDF) server through a polling procedure. The VSDF server then processes these data in order to extract observations (measurements or plots). Observations contain information about the estimated position and size of targets and the date and time of detections. These observations are sent to the tracker of the system for further processing. Ground coordinates corresponding to each fused observation are obtained using a calibration procedure, which is performed as a pre-processing step. It is assumed that the 3-D structure captured by each camera can be modelled as a plane, which is approximately true for most airport (and even road) applications.

3 BACKGROUND EXTRACTION TECHNIQUES AND POLYGONS

For the reliable detection of moving targets in the field of view of each camera, the estimation of the background and its periodic update are required. This is a very demanding problem especially for outdoor environments, where external factors such as camera oscillations, weather, gradual or sudden illumination changes and/or movements of objects belonging to the background affect the detection of moving targets. Such problems are usually addressed either by techniques aiming to update automatically the background (Gupte et al, 2002) (e.g. by taking a weighted average of the current background and the current frame of the video sequence) or by complex techniques (Borg et al, 2005), which apply statistical models for the estimation of each pixel value (e.g. mixture of Gaussians (Stauffer and Grimson, 1999), colour and edge fusion method (Jabri et al, 2000) etc).

In this paper, four state-of-the-art background modelling, subtraction and update techniques were extended so that they are applied only within specific regions of interest, defined by a set of

polygon-shaped detectors. The four methods are the Bayes technique (Li et al, 2003), the mixture of Gaussians (KaewTraKulPong and Bowden, 2001), the reliable background subtraction and update (Lluis et al, 2005) and finally, the non-parametric model for background subtraction (Elgammal et al, 2000). These extensions are seen to result to a very significant reduction of the complexity and execution times, as demonstrated in the experimental results section. Therefore, even techniques with increased computational complexity, like the Bayes-based or the Non-parametric Model approaches can be considered suitable for integration in real-time systems using the proposed technique. A sample foreground mask for each of the four methods is shown in Figure 1.

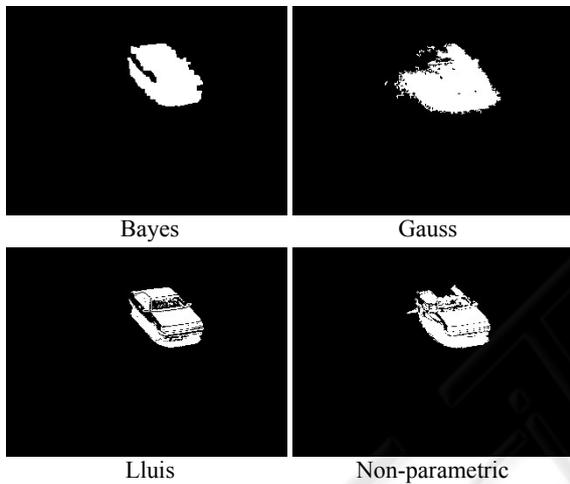


Figure 1: Foreground masks for the four methods.

Autoscope's concept of limiting the application of background extraction techniques to rectangular areas is an effective way of reducing execution time. However, it also limits the ability of the user to design more efficient detector configurations. To solve this problem, general polygon-shaped VDs were supported in the proposed approach. Furthermore, an additional "Sensitivity Indicator" property was added to each VD, in order to make the motion detection system even more flexible.

An off-line tool called "PolyMapper" was implemented to allow the user to define polygons of any shape and size depending on the scene structure and their specific needs and to adjust a threshold (sensitivity indicator) for each polygon, giving the percentage of pixels that have to be part of the foreground for the sensor to be considered as active. PolyMapper was built using the Qt library and can run both under Windows and Linux.

A sample frame and mask with the polygonal sensors marked is shown in Figure 2. When the percentage of foreground pixels is over a threshold, the sensor is highlighted.

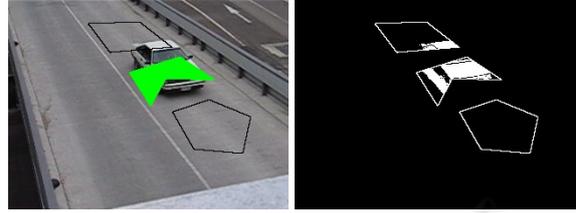


Figure 2: Sample frame and mask with polygonal sensors.

4 RESULTS AND CONCLUSIONS

For the purposes of this paper, the use of the new sensor was tested on traffic sequences with three different resolutions (320x740px, 640x480px, 768x576px). To have a more quantitative view of the time gain achieved with the polygon sensors, the four methods were applied on 50 frames of all three sequences and the frame rates achieved can be seen in Figure 3 for both the entire picture and the polygonal areas. These results do not include the time of the frame capturing process. For these tests, the methods were implemented using C++ and the OpenCV library. The system used was an Intel Pentium 4 3.2GHz with 1GB of RAM running on Windows XP Pro.

Finally, Figure 4 illustrates the percentage of decrease in execution times accomplished for the specific sequences.

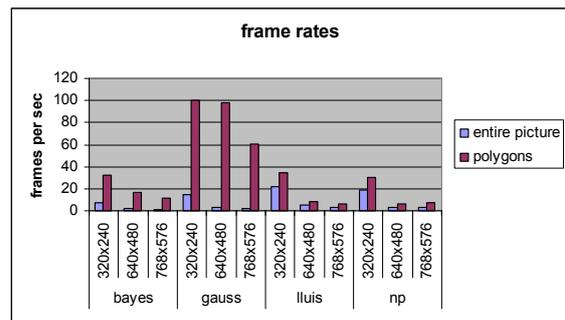


Figure 3: Test results, chart of frame rates.

The use of polygonal sensors to monitor traffic is proved to be notably effective. The execution times of modern but time consuming algorithms were decreased, allowing for use in real time applications.



Figure 4: Test results, chart of decrease in execution times.

Moreover, the polygonal shape gives flexibility to monitor areas that could not be covered with orthogonal sensors and the sensitivity indicator provides a way to parameterize each sensor separately, according to the user needs.

The performance of the four background extraction methods was also evaluated. The Bayes method, although it benefits from the proposed technique, does not provide satisfactory results in cases of slowly moving targets and it still remains quite slow. The Gauss method is faster but is not suitable for outdoor scenes, since it has problems coping with shadows. Results from the Lluis method deteriorate as the sequence resolution is increased. Finally, the non-parametric model method which provides the best foreground masks, benefits a lot from this technique, thus, it can be considered for real time applications.

In general, the obtained results are very promising and show great potential for the new sensor to be integrated as an alternative that can replace the Autoscope sensor for target tracking applications, similar to those developed by INTERVUSE project. Hardware implementations of the new algorithms may further reduce the computational costs and allow for the production of embedded systems such as Autoscope.

ACKNOWLEDGEMENTS

This work was supported by the General Secretariat of Research and Technology Hellas under the InfoSoc "TRAVIS: Traffic VISual monitoring" project and the EC under the FP6 IST Network of Excellence: "3DTV-Integrated Three-Dimensional Television - Capture, Transmission, and Display" (contract FP6-511568).

REFERENCES

- ICAO Document, 1986. 9476-AN/927: Manual of Surface Movement, Guidance, and Control Systems
- Besada, J.A., Garcia, J., Portillo, J., Molina, J.M., Varona, A., Gonzalez, G., 2005. Airport surface surveillance based on video images, *IEEE Transactions on Aerospace and Electronic Systems*, 41 (3), 1075 – 1082.
- Pavlidou, N., Grammalidis, N., Dimitropoulos, K., Simitopoulos, D., Strintzis, M.G., Gilbert, A., Piazza, E. Herrlich, C. and Heidger, R., 2005. Using intelligent digital cameras to monitor aerodrome surface traffic, *IEEE Intelligent Systems Magazine*, Vol. 20, No. 3, pp. 76-81.
- Image Sensing System, <http://www.imagesensing.com>
- Michalopoulos, P. G., Jacobson, R. D. Anderson, C. A. and DeBrucker, T. B., 1993. Automatic Incident Detection Through Video Image Processing, *Traffic Engineering and Control*. Vol. 34, No. 2, pp. 66-75.
- Michalopoulos, P. G., 1991. Vehicle Detection Video Through Image Processing: The Autoscope System, *IEEE Trans. on Vehicular Technology*, Vol. 40, No. 1.
- Gupte, S., Masoud, O., Martin, R.F.K. and Papanikolopoulos, N.P., 2002. Detection and Classification of Vehicles. In *IEEE Transactions on Intelligent Transportation Systems*. Vol. 3, No. 1.
- Borg, M., Thirde, D., Ferryman, J., Fusier, F., Valentin, V, Brémond, F. and Thonnat, M., 2005. Video Event Recognition for Aircraft Activity Monitoring. In *The 8th International IEEE Conference on Intelligent Transportation Systems*, Vienna, Austria.
- Stauffer C. and Grimson, W.E.L. 1999. Adaptive background mixture models for real-time tracking, In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 246-252.
- Jabri, S., Duric, Z., Wechsler, H., Rosenfeld, A., 2000. Detection and location of people in video images using adaptive fusion of color and edge information. In *15th International Conference on Pattern Recognition*.
- Liyuan Li, Weimin Huang, Irene Y.H. Gu and Qi Tian, 2003. Foreground Object Detection from Videos Containing Complex Background. In *International Multimedia Conference*.
- Elgammal, A, Harwood, D. and Davis, L., 2000. Non-parametric Model for Background Subtraction. *Computer Vision – ECCV 2000*.
- KaewTraKulPong, P., Bowden, R., 2001. An Improved Adaptive Background Mixture Model for Real-time Tracking with Shadow Detection. In *Proc. 2nd European Workshop on Advanced Video Based Surveillance Systems*, AVBS01.
- Lluis, J. Miralles X. and Bastidas, O., 2005. Reliable Real-Time Foreground Detection for Video Surveillance Applications. In *VSSN'05*.