Real-time target tracking using a Pan and Tilt platform

Moulay A. Akhloufi

Abstract—In recent years, we see an increase of interest for efficient tracking systems in surveillance applications. Many of the proposed techniques are designed for static cameras environments. When the camera is moving, tracking moving objects become more difficult and many techniques fail to detect and track the desired targets. The problem becomes more complex when we want to track a specific object in real-time using a moving Pan and Tilt camera system to keep the target within the image. This type of tracking is of high importance in surveillance applications. When a target is detected at a certain zone, the possibility of automatically tracking it continuously and keeping it within the image until action is taken is very important for security personnel working in very sensitive sites.

This work presents a real-time tracking system permitting the detection and continuous tracking of targets using a Pan and Tilt camera platform. A novel and efficient approach for dealing with occlusions is presented. Also a new intelligent forget factor is introduced in order to take into account target shape variations and avoid learning non desired objects. Tests conducted in outdoor operational scenarios show the efficiency and robustness of the proposed approach.

Keywords—Tracking, surveillance, target detection, Pan and tilt.

I. INTRODUCTION

TARGET tracking is an important field that attracted a lot of interest in computer vision. Available techniques for object tracking use different approaches like background substraction or modeling, particle and Kalman filtering, segmentation, supervised learning, etc. [1]

The vast majority of research in tracking deals with image sequences captured by a static camera. The recent increase of interest for studying the tracking techniques is mostly driven by surveillance and security applications, where traditionally the moving cameras were monitored by security personnel. However, with today increasing treats we need more cameras, thus making human monitoring difficult. Developing new algorithms for automatically tracking detected targets become necessary. In the case of moving target captured by a moving camera the challenge become more important. Recently, we see an increase of interest in dealing with these cases [2]-[4]. Interesting work has been done in the past in the area of

M. A. Akhloufi is with Centre de Robotique et Vision Industrielles, Levis, QC G6V 6Z9 Canada and with Vision Systems Laboratory, Laval University, Quebec, QC G1V0A6 (phone: 418-833-1965 x3275; fax: 418-833-8726; e-mail: moulay.akhloufi@crvi.ca).

detecting and tracking objects in airborne video imagery where the targets and the camera are both moving [5], [6]. More recently, we developed an efficient approach for detecting and tracking small moving objects present in low resolution and noisy images [7]. The work presented in [7] dealt with images coming from static cameras and airborne video images captured by a UAV. However, little work deal with tracking moving targets using a Pan and Tilt platform [8], [9]. Most of the work presented in the past does not deal efficiently with occlusions and target shape variations. In this work we propose a new approach aiming to solve these problems.

II. PROPOSED APPROACH

In this work, we present a real-time target tracking approach using and Pan and Tilt camera system. The proposed approach permits the continuous tracking of targets and handles efficiently occlusions due to objects present in the scene. The following steps are used for tracking:

1) Target detection and learning in the detection zone;
2) Appearance based tracking:
   a. Particle filtering;
   b. SSD and PCA learning/recognition;
3) Occlusion handling;
4) Pan and Tilt control.

The following sections give the details about these different steps.

III. TARGET DETECTION

In order to detect the target, an approach based on spatial and temporal discreet wavelets processing developed in a previous work is used. This algorithm is very efficient and robust to noisy images. Detailed description of this approach is given in [7].

In order to start tracking, an alert zone is defined and the target detection algorithm starts processing this area. When a target is detected the information about its position is sent to a target extraction algorithm. In this step N successive frames are used in order to extract the shape of the object based on the detected motion (Fig. 1 and Fig. 2). This extracted object is then learned for PCA recognition and the particle filter takes in charge the tracking of this target using the Pan and Tilt camera system as described below.

The target detection algorithm permits the detection and tracking of multiple targets in an image [7]. However, in this

work we start tracking the first object entering the alert zone since only one Pan and Tilt system is used and the camera is moved. It cannot track multiple objects heading in different directions. The extension to multiple targets using multiple Pan and Tilt systems is straight forward.

Fig. 1 Detection of a vehicle entering the target detection zone (the first learned shape used in subsequent tracking).

Fig. 2 Pedestrian detection and shape learning.

IV. PARTICLE FILTER

Particle filters have been widely used in visual tracking. These filters are more general than Kalman filters who have restrictive linear Gaussian assumptions regarding the transition models [10]. The particle filter is a sequential Monte Carlo algorithm, a sampling method for approximating a distribution that makes use of its temporal structure [11]. With this filter, we can predict the position of the object based on its previous state. At a high frame rate, the displacement of the target and its position at *t+1* is limited to areas around its position in the previous frame at *t*. Also the change in orientation is limited to small variations. In this work a particle filter is used to create randomly search zones around the area of the detected object at instant *t*. It uses a Gaussian distribution with a certain standard deviation for each parameter of the search window. The search window is characterized by the following parameters: position, height/width ratio, scale (in comparison with a fixed reference window), rotation angle and skew.

Fig. 3 shows an example of search windows created using a particle filter based on the current detected target.

V. LEARNING AND RECOGNITION

In this work, we use an appearance model approach for effectively tracking the previous selected object. A model of the selected target is built using two approaches described below.
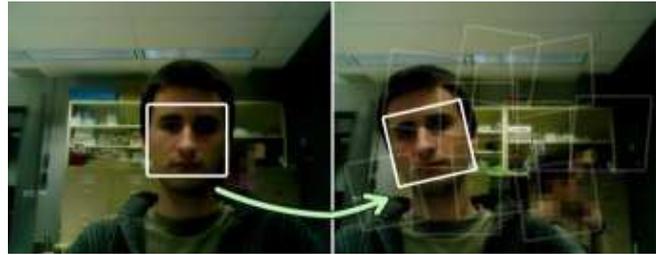

Fig. 3 Particle filter search windows

The first approach consists of computing a sum of squared difference (SSD) between the search window and the mean target image (Fig. 4). When the target is detected, the mean image is updated with the new target image. This permits to take into account object variations over time.

Fig. 4 Detected target in the image (upper image) and mean target image (lower image)

The second approach uses a principal component analysis approach (PCA) for learning and recognition. PCA is a very well known approach for projecting the object in the best eigenspace representing all the learned objects (Fig. 5). It was largely use in face recognition and is known as eigenfaces technique. This approach is used for object recognition. In this work, the PCA is used incrementally on the selected target. Once the target is recognized, it is learned and added to the eigenobjects. The previous computed mean image is used in the learning process.

Fig. 5 Example of the first eigenobjects (eigenfaces)

However since the object can move in different directions and its shape can change, we introduce an intelligent forget factor.

When the computed SSD and PCA distance are within a certain thresholds, the mean image is reset. Also, only the last N detected images are used to compute the new mean target image and in the eigenobjects learning phase as in (1).

This step take into account the results of the occlusion detection described below before updating the mean image, in order to avoid the use of a non desired object in the learning and recognition phases.

$$\begin{cases} C_1 = d_{SSD} > T_{SSD}^1 \ \& \ d_{PCA} > T_{PCA}^1 \\ C_2 = d_{SSD} > T_{SSD}^2 \ \& \ d_{PCA} > T_{PCA}^2 \end{cases}$$

$$\left\{ \forall t > t_N : C_1 \ \& \ C_2 \rightarrow \overline{M} = \sum_{k=t-N}^{t} I_k \right\} \tag{1}$$

Where:

d_{SSD} is the SSD computed distance.

d_{PCA} is the PCA computed distance.

T_{SSD}^i is the SSD threshold.

T_{SSD}^i is the PCA threshold.

$\overline{M}$ is the mean target image.

VI. OCCLUSION HANDLING

Occlusion is a big problem when tracking selected targets. A target can be lost when occluded and tracking can start with a different object. In order to avoid this problem, we developed a novel approach for handling occlusions in this kind of selective tracking.

When the SSD is above a certain threshold for N successive frames, the object is considered lost and an extended search is performed. The number of particles and the standard deviation of the particle filter are increased. The search zone is increased around the last target detection position.

Based on the previous state of the target the search zone is increased in a non symmetrical way, thus giving higher weights to the areas where the probability P of finding the object is higher.

VII. PAN AND TILT SYSTEM CONTROL

In this work, we need a fast Pan and Tilt system in order to track in real-time fast moving objects like cars. We used an industrial Pan and Tilt system from Directed Perception [12]. The model is a PTU-D46-17 characterized by its high speed: up to 300°/sec and high resolution: 0.01° (Fig. 6).

Fig. 6 Pan and Tilt unit PTU-D46-17with a 1024x768 camera

The Pan and Tilt system is moved using the extracted information from the particle filter target tracking. When the target is identified as a positive detection, its position is sent to an algorithm for the estimation of the new Pan and Tilt position. The visual servoing loop is described in the diagram of Fig. 7.
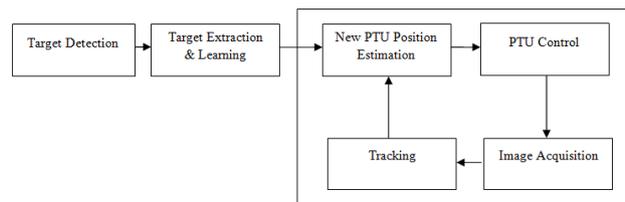

Fig. 7 Diagram of the visual servoing loop

VIII. EXPERIMENTAL RESULTS

Experiments were conducted in indoor and outdoor environment.

An example of indoor tests is showed in images of Fig. 8. The tracking of the face is conducted in real-time (25 frames/second). The images show the efficient handling of occlusions. The face is tracked even if it is occluded in multiple frames.

Fig. 9 to Fig. 11 show some results in an outdoor environment. We can see that the tracking is fast and can track

in real-time vehicles as shown in Fig. 9. Also, the results show that the proposed system is robust to outdoor conditions (Fig. 9-Fig. 11) and can handle occlusions efficiently (Fig. 10).

Tracking in outdoor scenarios was done at a frame rate of 20 frames/second in a Firewire 1394b camera with 1024x768 pixels resolution. The program is running in a PC with a Quad Core Q8300 2.5Ghz and 2 GB memory.

In order to obtain real-time performance, the algorithms were developed in C++ and optimized for faster processing. Multi-Threading was used to balance the processing load. Each of the following algorithms was assigned to a thread:

- Target detection and learning in the detection zone;
- Target tracking (particle filter);
- Image acquisition;
- Pan and Tilt control.

An intelligent tracking manager algorithm was developed in order to optimally monitoring and dispatching the information extracted from the different modules.

The experimental tests show that the proposed technique is robust and permit to track efficiently and in real-time targets of interest.

IX. CONCLUSION

In this work we presented a real-time tracking system for detecting and tracking targets of interest using a Pan and Tilt camera platform. In the proposed approach we use a particle filter for tracking the selected target.

A new intelligent forget factor is introduced. It permits to take into account target shape variations and avoid learning non desired objects during the appearance based tracking step.

A novel and efficient approach for dealing with occlusions is also presented.

Experiments conducted in indoor and outdoor environments show the high performance achieved by the proposed system. Real-time tracking was done at a rate of 25 frames/second for indoor and 20 frames/second for outdoor tests.

Future work includes increasing the performance of the system by performing the tracking in a scale space framework.

ACKNOWLEDGMENT

REFERENCES

[1] Yilmaz, A., Javed, O., and Shah, M., "Object Tracking: A Survey," ACM Journal of Computing Surveys, 38 (4), 1-45 (2006).
[2] Shah, M., Hakeem, A., and Basharat, A., "Detection and Tracking of Objects from Multiple Airborne Cameras," SPIE Newsroom, (2006).
[3] Kang, J., Cohen, I., Medioni, G., and Yuan., C., "Detection and Tracking of Moving Objects from a Moving Platform in Presence of Strong Parallax," Proc. IEEE Int. Conf. on Comp. Vision, (2005).
[4] Yu, Q., Cohen, I., Medioni, G., Wu, B., "Boosted Markov Chain Monte Carlo Data Association for Multiple Targets Detection and Tracking," Proc. Int. Conf. on Patt. Recog., (2006).
[5] Cohen, I. and Medioni, G., "Detecting and Tracking of Objects in Airborne Video Imagery," Proc of IEEE Compt. Vision and Patt. Recog. Interpretation of Visual Motion Workshop, (1998).
[6] Davies, D., Palmer, P. and Mirmehdi, M., "Detection and Tracking of Very Small Low Contrast Objects," Proc. British Machine Vision Conference, 599-608 (1998).
[7] Akhloufi, M., "Tracking multiple-sized objects in low resolution and noisy images," Airborne Intelligence, Surveillance, Reconnaissance (ISR) Systems and Applications VI. Proceedings of the SPIE, Volume 7307, pp. 73070J-73070J-12 (2009).
[8] Shah, H., and Morrell, D., "An adaptive zoom algorithm for tracking targets using pan-tilt-zoom cameras," Proceedings of the 2004 IEEE International Conference on Acoustics, Speech, and Signal Processing, Montreal, Canada (2004).
[9] Cindy, X. Collange. F. Jurie, F. Martinet, P. "Object tracking with a pan-tilt-zoom camera: application to car driving assistance", Proceedings of IEEE International Conference on Robotics and Automation. pp. 1653- 1658 vol.2, Seoul, Korea (2001)
[10] Cuevas, E., Zaldivar, D., and Rojas, R., "Kalman filter for vision tracking," Freie Universität Berlin Institut für Informatik Technical Report B 05-12, (2005).
[11] Cuevas, E., Zaldivar, D., and Rojas, R., "Particle filter in vision tracking," Freie Universität Berlin Institut für Informatik Technical Report B 05-13, (2005).
[12] Directed Perception PTU, http://www.dperception.com/

Fig. 8 Results of face tracking under occlusion

Fig. 9 Results of vehicle tracking

Fig. 10 Results of pedestrian tracking and occlusion handling
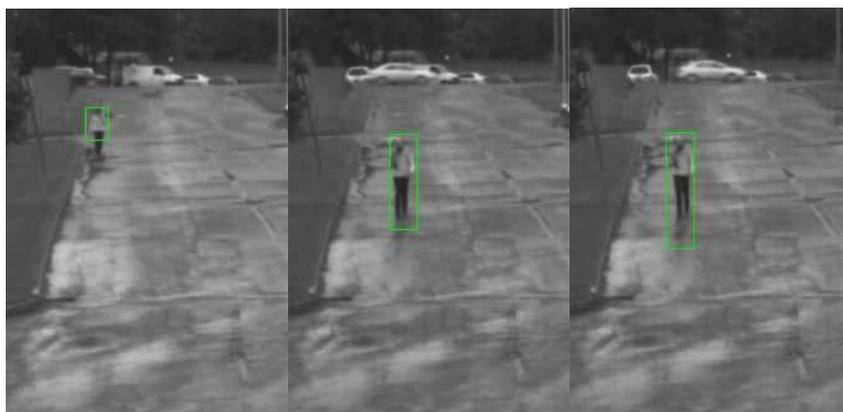
Fig. 11 Results of tracking a pedestrian in rainy conditions