Object tracking with UAVs

Arnaldo Abrantes
Instituto Politécnico de Lisboa

Pedro Mendes Jorge
Instituto Politécnico de Lisboa

Object tracking with UAVs

João Palma [1,2]
Arnaldo Abrantes [1]
Pedro Mendes Jorge [1]

[1] Department of Electronics and Telecommunications and Computers, ISEL
[2] TEKEVER Information Technology

Abstract

This paper presents an automatic object tracking system that processes video captured by a camera installed in a UAV. The system returns the position of a given target which can be integrated with the GPS navigation control system of the aerial vehicle.

1 Introduction

The use of UAVs is growing up at an impressive speed. Every day we heard about military operations concerning UAVs, not only for vigilance purposes, but also for operations involving the attack of predefined targets. [4]
A large number of projects have been presented in the last few years, related with tracking using UAVs. COCOA [2] and MODAT [3] are two examples of systems that perform real time tracking of objects captured by a video camera installed on UAVs. The implementation of both of these systems adopts an architecture where the output (estimation of the target position) is obtained by applying to the input (captured video) a chain of three consecutive standard image analysis techniques: Image Registration, Image Segmentation and Object Tracking. According to MODAT, this system has a success rate of 80%. However, the authors refer also a false positive rate of 18%.
In this work, it was developed an automatic object tracking system that processes at real time the video captured by a camera installed at UAV. The system returns the position of the desired target (an object that is moving on the field of view of the camera) and can be integrated with the GPS navigation control system of the aerial vehicle.
The proposed system is based on the Mean shift algorithm to perform object tracking. In order to improve the results, this standard tracking technique was enhanced by previously filtering a back projected image (used as the input of Mean shift), using the target predicted position given by a Kalman filter.

2 Object tracking with UAVs

The proposed system is described by the block diagram presented in Figure 1. The system contains three main modules, which are summarized in the following subsections.

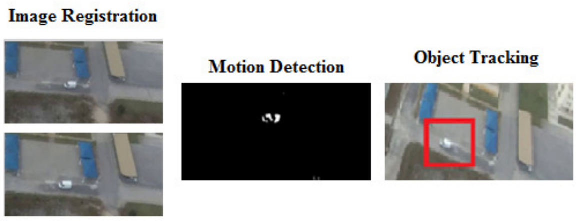


Figure 1: System architecture with its three modules: Image Registration, Motion Detection and Object Tracking

2.1 Image Registration

This first module is responsible to align the captured frames by incrementally making a mosaicing, where all the frames belong then to a unique world referential. This is a crucial operation to make possible to detect object motion with image subtraction. Image registration is accomplished by combining the following techniques: camera calibration, image homography, feature extraction, and optical flow estimation.
A camera calibration method [1] is used to correct geometric deformations present in the video frames. In this kind of projects, these deformations are usually strong, as the used lenses have a large field of view, which generate significant radial and tangential distortions (Fish-Eye Lens). The align phase uses feature extraction to detect interest points (corners) and optical flow to follow those features in the next frame. The set of corresponding points is then used to estimate a homography matrix, which is finally used to align consecutive frames into the same geometric referential.

2.2 Motion Detection

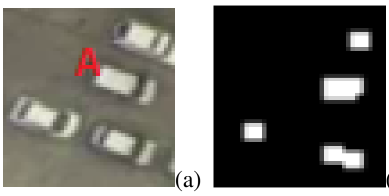
Motion detection uses two different techniques: image subtraction and color information.
In the first one, the aligned frames are returned from the previous module and subtracted to obtain a difference image that can be used as a starting point to detect the moving objects.
A second approach is the use of a color histogram to model the desired target, and subsequently compute, for each new frame, the probability of each pixel color was produced by the moving target. This new image, containing the referred probabilities, is usually designed as back projected image.

2.3 Object Tracking

The system tracks one of the objects returned by the motion detection module: that one selected by the user.
Tracking is then accomplished by combining Meanshift and Kalman filter algorithms.
Meanshift is used to estimate the object location based on the local maximums (modes) of the back projected image (i.e., a probability density function). The algorithm performs an iterative local optimization, searching in each iteration for the maximum of the pdf in a small square search window, centered at the previous estimation.
Kalman Filter is an algorithm that produces estimates of state variables based on a series of measurements (observations). In this model, it is assumed that prediction and measurement values are affected by white Gaussian noise with zero mean. A dynamic model with constant velocity is used in this work.
The tracking algorithm operates as following: first, the target's location is predicted using the Kalman filter prediction equation, then Meanshift is used to refine the predicted position, which is finally used as the observation in the last step of the Kalman filter.

2.3.1 Filtering back projected image

In several tracking problems, there are many small potential targets (each one with only a few numbers of pixels) and most of them indistinguishable in terms of color (see Figure 2.a). In this case, color information (e.g. color histograms) is insufficient to detect the target as can be seen by observing the high noisy back projected image (Figure 2.b). Figure 2.a is an example of the back projected image from a window near the target location of this sequence. All the vehicles in this sequence exhibit the same color and, based only in this information, it is very difficult to estimate the correct target location.



(a) (b)

Figure 2: Binary representation of back projected image (a), based on the color histogram of car A (a)

The existence of several local maximum within the Meanshift search window can become a point of failure when applied to object tracking.

To alleviate this difficulty a Gaussian filter, centered on the region identified as the predicted target location, is applied to the back projected image. This filtering operation is able to reduce the noise in the back projected image and therefore improving the results of Meanshift (Figure 3.b).

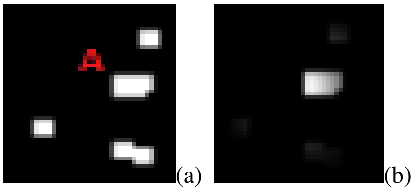


Figure 3: Application of a Gaussian filter to the back projected image of Figure 2.a. Gaussian is centered at point A, predicted by the Kalman filter

Therefore, the block diagram of Figure 4 describes the tracking methodology of this work.

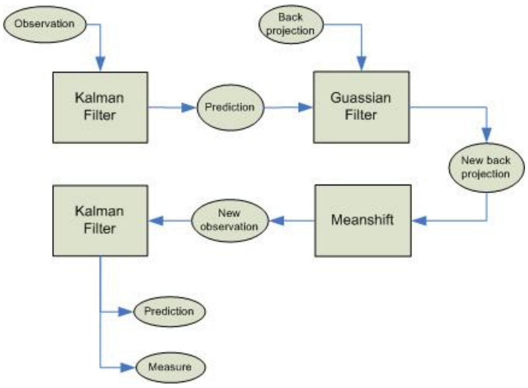


Figure 4: Back projection's correction steps

If occurs an error in object detection caused by Meanshift, Kalman filter can lose the object of interest. Correcting the back projected image is an approach to reduce the processing errors of Meanshift. This technique generates a new back projected image (Figure 6.b) by applying a Gaussian filter to the previous back projected image (Figure 5.b). Once the Gaussian's center is fixed at Kalman's predicted point (Figure 5.a), the closest regions will have a much lower probability, increasing Meanshift results (Figure 6.a)

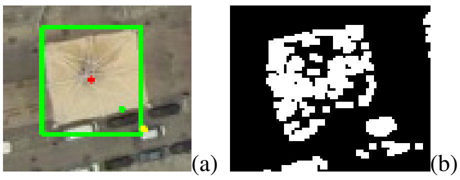


Figure 5: Kalman and Meanshift results (a), based on the original back projected image (b): Observation (red), Correction (green), Prediction (yellow) and Meanshift result window (green rectangle)

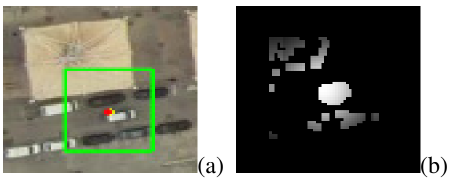


Figure 6: Kalman and Meanshift results (a), using the corrected back projected image (b)

3 Results

The image sequences used in the experimental tests presented in this work were acquired by the UAV shown in Figure 7, developed by TEKEVER Autonomous Systems.

Figure 7: UAV (AR4 Light Ray) used during the development of this work

Figure 8 shows the result with the tracking algorithm without applying the Gaussian Filter. This images shows a failure of tracking.
The results of the same sequence with the filtered back projected image are shown in Figure 9. Now, the system is able to track the desired vehicle correctly.

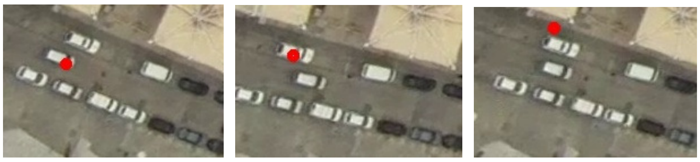


Figure 8: Tracking results without filtering back projection matrix

Figure 9: Tracking results with back projection matrix filtered

4 Conclusion and Future Work

This paper describes a real time tracking system for UAVs. The proposed method is based on the combination of Meanshift and Kalman algorithms, using as input back projection matrix obtained by measuring, at each frame, the difference between local color histogram and a model histogram representing the tracked target. The systems tends to fail whenever two similar (in color) targets are close, as in this case, Kalman filter may be driven by a false observation.
In order to alleviate this difficulty and enhancing the robustness of the method, the back projected image is previously processed, filtering out the data further from the predicted position of the target, before being used by Meanshift algorithm.
The results obtained with this systems show that it is capable of accurately tracking a small target even if other distracting targets are close. However, there are situations when the system loses its target and in that case it cannot recover from this failure. This is the main problem of the proposed method which can be ameliorated by extending it to a multimodal technique like, e.g., particle filters.
The target trajectory detected with this algorithm is still not used to control the UAV. The integration with the GPS navigation control system will be a next stage of this project.
Some future work on multi-target tracking should also be addressed, namely, to maintain multiple targets in the UAV field of view.

References

1. **Bouguet, J.Y.** Camera Calibration Toolbox for Matlab. [Online] http://www.vision.caltech.edu/bouguetj/calib_doc/.

2. **Shah, Saad Ali and Mubarak.** COCOA - Tracking in Aerial Imagery. *Airborne Intelligence, Surveillance, Reconnaissance (ISR) Systems and Applications III.* 2006, Vol. 6209.

3. **Ibrahim, Aryo Winam, et al.** Moving Objects Detection and Tracking Framework for UAV-based Surveillance. *Fourth Pacific-Rim Symposium on Image and Video Technology.* 2010.

4. US Drone Strikes Kill Seven Alleged Al-Qaida Yemen. *The Guardian.* [Online] http://www.theguardian.com/world/2013/aug/07/us-drone-strikes-kill-seven-alleged-al-qaida-yemen