# ON-BOARD UAV VIDEO PROCESSING FOR GROUND TARGET TRACKING

**Cristian VIDAN, Gavril ALEXANDRU, Razvan MIHAI, Florin CATARGIU**

Military Technical Academy, Bucharest, Romania (cristian.vidan@mta.ro)

DOI: 10.19062/2247-3173.2018.20.36

***Abstract:*** *This paper presents a target tracking system based on video processing on the UAV. The primary role of the system is to overlay the center of the target selected by the ground operator with the center of the image displayed on the GCS (Ground Control Station) monitor. In fact, the coordinates of the two-dimensional object are initially determined in a Cartesian system, and later on, based on these coordinates are determined angular deviations in azimuth and elevation. Once the angular deviations have been obtained, the transmission to the gimbal controller on which the optical sensor is mounted will be done in digital format. This makes the target always be at the center of the image, which makes the image operator work significantly easier.*

***Keywords:*** *target tracking, video processing, on-board, UAV, software.*

## 1. INTRODUCTION

As we know from the definition, target tracking is the prediction of the future location of a dynamical system based on its estimates and measurements.

Object detection and target tracking are two important domains of computer vision that if connected to artificial intelligence (AI) can make robotic systems more accurate and useful to humans. Also, a challenging problem is the strong weight and area constraint of embedded hardware that limits the drones to run computation intensive algorithms, such as deep learning, with limited hardware resource.

There is a wide range of applications when it comes to tracking the various objects of interest, hence the motivation of researchers worldwide to focus their attention on this area. Video surveillance is one of the most important. Surveillance systems not only record visual information but also extract data on motion, and more recently analyze suspicious changes within the frame. With these surveillance systems can be visually monitored other planes, vehicles, animals, micro-organisms or other moving objects but detecting and tracking people is of great interest. For example, human counting applications can provide important information about public transport, traffic congestion, trade and security.

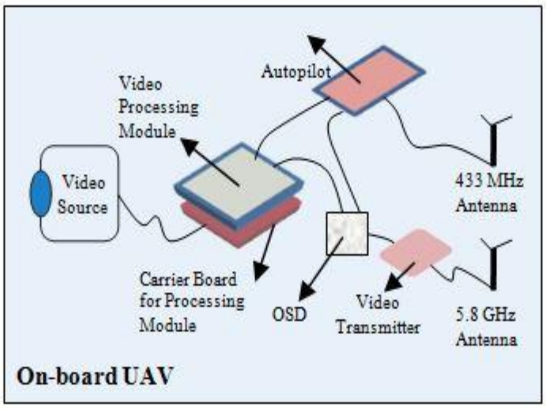
## 2. SYSTEM DESCRIPTION

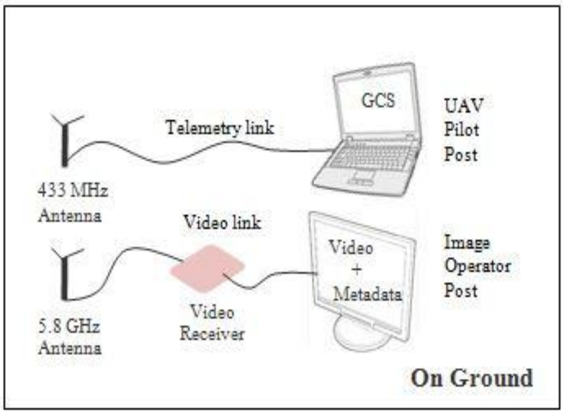
In order to massively deploy drones and further reduce their costs, it is necessary to power drones with smart computer vision and autopilot. This increases the accuracy of detecting and tracking the target and the ability to accomplish specific missions.

The objective of video tracking is to associate target objects in consecutive video frames.

The association can be especially difficult when the objects are moving fast relative to the frame rate. Another situation that increases the complexity of the problem is when the tracked object changes orientation over time. For these situations video tracking systems usually employ a motion model which describes how the image of the target might change for different possible motions of the object.

The diagram of the target tracking system that was designed is shown in Fig. 1. It can be seen from the diagram that the UAV's main component and the brain of the system is Nvidia Jetson TX2 video signal processing module. Jetson TX2 is very fast and most power efficient (7.5 W) if you take into account its dimensions. It has 8 Gb of memory and almost 60 Gb/s of memory bandwidth. Also, it features a variety of standard hardware interfaces that make it easy to integrate with a wide range of carrier boards.




a. On-board UAV system hardware　　b. On Ground UAV system hardware

**FIG.1** Target tracking system diagram

The tracking target system uses two data links: the telemetry channel that is set to the 433 Mhz frequency and provides command and control of both the autopilot and the optical sensor and the video channel that is set to the 5.8 Ghz and provides download link to the video receiver to be displayed on the image operator post monitor.

## 3. SOFTWARE ARCHITECTURE

Tracking algorithms can be classified into two major groups, namely state-space approach and kernel based approach. State-space approaches are based largely on probability, stochastic processes and estimation theory, which, when combined with systems theory and combinatorial optimization, lead to a plethora of approaches, such as Kalman filter, Extended Kalman Filter (EKF) [1], Unscented Kalman Filter (UKF) [2], Particle Filter (PF) [3]. The ability to recover from lost tracks makes state-space approach one of the most used tracking algorithms. However, some of them require high computational costs so they are not appropriate for real time video surveillance systems.

The Mean Shift (MS) algorithm is a non-parametric method which belongs to the second group. MS is an iterative kernel-based deterministic procedure which converges to a local maximum of the measurement function under certain assumptions about the kernel behaviors [4]. CamShift (Continuously Adaptive Mean Shift) algorithm [5] is based on an adaptation of mean shift that, given a probability density image, finds the mean (mode) of the distribution by iterating in the direction of maximum increase in probability density. CamShift algorithm has recently gained significant attention as an efficient and robust method for visual tracking. A number of attempts have been made to achieve robust, high-performance target tracking.

The Camshift principle is described as follows. The first step is to convert the image into color probability distributions, depending on the object's color histogram. The second step is the initialization of the dimension and the position of the search window. The last step is the adjustment with the result obtained in the previous step. After completing the 3 steps, we can locate the center of the object in the current image.

Firstly, we calculate the color histogram of the object. The RGB color space has weakness in representing shading effects or rapid illumination changing. To solve this problem, the Camshift algorithm adopts the HSV [6] ( Hue Saturation Value) color space to describe the object. For each frame, the chrominance histogram is calculated as the following: Suppose that the number of total pixels in the frame is **n** , the chrominance grade of the histogram is **m** , and the chrominance grade index corresponding to the i-th pixel $x_i$ is $c(x_i)_{i=1,\ldots,n}$ ,then the element of histogram $\{q_u\}_{u=1,\ldots,m}$ of this frame is

$$q_u = \sum_{i=1}^{n} \delta\left[c(x_i) - u\right], u = 1, 2, \ldots, m \tag{1}$$

where $\delta$ is unit impulse function. The formula of histogram normalization is as following:

$$\left\{ p_u = \min\left( \frac{255}{\max(q)} q_u, 255 \right) \right\}_{u=1,\ldots,m} \tag{2}$$

We can normalize the value range of the histogram corresponding to every chrominance grade to the interval of [0, 255] so that through back projection the value of each pixel in the image is related to the corresponding to the value of each chrominance grade in the histogram. Thus the probability distribution graph of object-color is obtained.

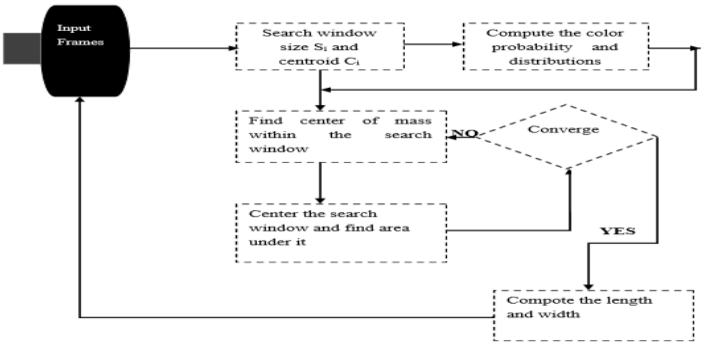



**FIG.2** Traditional CamShift Algorithm

## 4. RESULTS

When the video signal is transmitted to the ground, the operator selects the target of interest with four points. The next step is to frame the target in a red rectangle and to determine the coordinates of the center, so that the video camera ghimbal then makes the corrections necessary to overlay the center of the image and the center of target.

**a**) Selecting a frame    **b**) Selecting the target

**C)** Positioning the target at the center of the image

## 5. CONCLUSION

The algorithm has been tested under various conditions and the results are satisfactory. The target is detected and tracked without interruption, and the algorithm used is not influenced by the speed of the target.

## 6. FUTURE WORK

As future work we intend to improve performance and reduce costs by combining video tracking and image stabilization in a single device. Therefore, we intend to design a compact gimbal which is capable to detect and track multiple targets in the same time with a lower power consumption video stream processor. Also, artificial intelligence has developed quite a lot lately and it is increasingly such that on the future gimbal we will try to implement machine learning or deep learning algorithms that will greatly help both the pilot and the image operator to accomplish their critical missions.

## REFERENCES

[1] Yaakov Bar-Shalom and Thomas E. Fortmann. 1988, Tracking and Data Association. Academic Press, 1988.
[2] Simon J. Julier. and Jeffrey K. 1997 “A new extension of the Kalman filter to nonlinear systems,” in Proc. SPIE ,Vol. 3068, p. 182-193
[3] K. Nummiaro, E. Koller-Meier, and L. Van Gool, “A color-based particle filter,” in Proc.of the 1st Workshop on Generative-Model-Based Vision, June 2002, pp. 53–60.
[4] R. T. Collins. Mean-shift blob tracking through scale space. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2003.
[5] G. R. Bradski. “Computer vision face tracking for use in a perceptual user interface”, Intel Technology Journal, 2nd Quarter, 1998
[6] Nouar, O.-D.; Ali, G.; Raphael, C.2006,Improved Object Tracking With Camshift Algorithm, IEEE ICASSP 2006, Volume 2, 14-19 May 2006 .