Vision-based autonomous object tracking using multi-rotor UAV

Talha Farooq Hashmi,, Abdullah Rashid , Waleed Khan, Benish Sharfeen Khan*

Department of Electrical Engineering, Institute of Space Technology, Islamabad, Pakistan
(talhafarooq4@gmail.com, abdullah.rashid@outlook.com, waleedk2245@gmail.com, benish.sharfeen@gmail.com)

**Corresponding Author*

Abstract

In this paper- we are presenting a real-time method to detect and track an autonomous object using visual processing on an unmanned aerial vehicle using an on-board companion computer (Jetson-TX1) for image processing. The profile of objects, frame rate of images, and unexpected motion make it hard to detect and track the object for a long period. To cater to this, we came up with an algorithm that was developed for long-term tracking which makes use of Discriminative correlation filter with Channel and Spatial Reliability. The major restriction of our algorithm arises in the presence of occlusion, which was solved by creating a region of interest in the center of the frame in which the object will always reside. If the object exits the center region a command of left, right, top or bottom will be generated following the position of the object relative to our center position. These commands will be communicated to the UAV via Mavlink protocol. Experimental results show that we have achieved a long period of tracking with a good frame rate and eliminating spurious events and misdetections.

***Keywords**: Image processing, Drone, Companion computer, Onboard processing*

1. Introduction

With the rapid development in technology, we are looking towards drones as the future. Currently, common drone applications require dedicated human interaction, which is laborious, inefficient, restricted, and requires a skilled pilot. With our project, we are hoping to change this by integrating the intelligence within the drone making it fully autonomous and self-sufficient.

Nowadays, drones are being used in several areas such as photography, farming, environment protection, military, etc.; however, this project is intended for surveillance applications, specifically those involving object tracking.

The project can be broken into two work areas, hardware assembly/integration and the development of an object tracking algorithm. Hardware includes the custom-built drone fitted with an onboard microcomputer and a camera. The algorithm will run on the microcomputer, utilizing the camera attached with it, and in return, pass required flight commands to the flight controller. This operation is summed up in Fig 1.

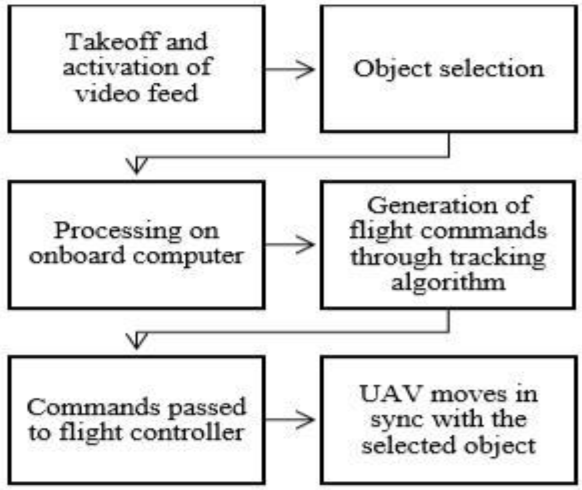

Figure 1: Architect of our system

The camera provides a video stream to the onboard computer (Jetson-Tx1), a tracking algorithm on the computer processes the video feed. It determines the position of the object to track on the frame and then generates flight motion commands which are then passed to Pixhawk to move UAV in sync with the object. This is a continuous process algorithm that sends commands and through feedback process drone syncs with object motion to maintain track additionally, we are also receiving data like inertial measurement unit (IMU) and battery status on ground station [1].

Figure 2: Drag an object to track

Initially, we came up with different image processing algorithms. Starting with HSV [2] discrimination base tracking but it was light-dependent, then SURF a detection algorithm this works with feature extraction and matching process on each frame but due to irregular motion of object and change of perspective made our search for an alternative [3], [4]. Then we integrate the Lucas-Kunade method (KLT) [5] to track motion and store the tracking coordinates but when several objects came in the video feed it gets disturbed. Moreover, the detection algorithm always starts detection from start. So, we need some good tracking algorithm which should detect in the n-1 frame and should keep up the track in the nth frame. [6] So, to overcome these issues we came up with a computationally efficient and robust algorithm, the discriminative correlation filter with Channel and Spatial Reliability which is explained in a further section. However, object tracking has a variety of uses, some of which are: surveillance and security, traffic monitoring, video communication, robot vision and animation.

Contribution

- A real-time method to detect and track an autonomous object using visual processing on an unmanned aerial vehicle (UAV) using an on-board companion computer (Jetson-TX1) for image processing.
- Robust discriminative correlation filter with Channel and Spatial Reliability algorithm is proposed for object tracking.

2. Hardware Assembly and Integration

2.1 Hardware Assembly

The nature of our project is such that we could not rely on premade multi-rotors or kits because we needed a large UAV that was capable of carrying at least 1.5Kg payload and could achieve longer flight times without any issues. Hence, we had to buy specific components from various vendors and manufacturers. Selection of UAV related components was done with the help of ECALC software. ECALC is an online software that simulates the flying characteristics and overall physical parameters of the UAV based on selected components [7]. It requires us to provide the basic configuration of UAV, frame specifications, battery specifications, atmospheric conditions in which the UAV is most likely to be flown, payload weight, propeller size and a few other parameters. The estimated performance parameters of our UAV in Fig 4. After the procurement of components, the UAV was assembled and basic flight test were successfully conducted before proceeding further [8].

Figure 3: Radio Telemetry Jetson Nano Pixhawk 2.1

Battery		**Motor @ optimum Efficiency**		**Motor @ Maximum**	
Load	12.51C	Current	6.59A	Current	15.63A
Voltage	21.22V	Voltage	21.63V	Voltage	21.16V
Rated Voltage	22.20V	Revolutions	6650rpm	Revolutions	6150rpm
Energy	111W/h	Electric Power	185.8W	Electric Power	330.7W
Total Capacity	5000mAh	Mech Power	163.3W	Mech Power	282.8W
Used Capacity	4250mAh	Efficiency	87.9%	Power Weight	617.9W/kg
Min Flight Time	4.1 min				280.3W
Max Flight Time	14.2 min			Efficiency	85.5%
Heavy Flight Time	23.6 min			Est.Temperature	41^{O}C
Weight	786 g				105^{O}F
	27.7oz			**Walt meter readings**	
				Current	62.52A
				Voltage	21.22V
				Power	1326.7W

Motor@hover		**Total Drive**		**Multicopier**	
Current	2.70A	Drive Weight	1.44/g	All-Up weight	2141g
Voltage	22.02V		50.8oz		75.5oz
Revolutions	31.01rpm	Thrust-Weight	3.1:1	Add Payload	3626g
Theoretic(Kg)	30%	Current@Hover	10.79A		127.9oz
Theoretic	47%	P(In)@Hover	202.7W	Max Tilt	45''
Electric Power	59.7UV	Efficiency@Hover	84.6%	Max.speed	46km/h
Mech Power	50.7W	Current@max	62.53A		26.6rmph
Power Weight	111.9UWkg	P(in)@rmax	1388.2W	Est.rate of climb	7.9m/s
	50.8UW	P(out)@max	1131.3W		1.565ft/min
Efficiency	85.3%	Efficiency@max	81.5%		
Est Temperature	28^{O}C			Total disc Area	45.60dm*
	82^{O}F				706.8in^2
Specific Thrust	9.0/g/w			With Rotor tail	
	0.320z/w				

Figure 4: Drag an object to track

2.2 Integration

The integration stage involves the communication between the flight controller and the onboard microcomputer. Restricted by the flight controller, this communication can only be done using Mavlink protocol. MAVLINK is a serial protocol most used to send data between vehicles and ground stations [9].

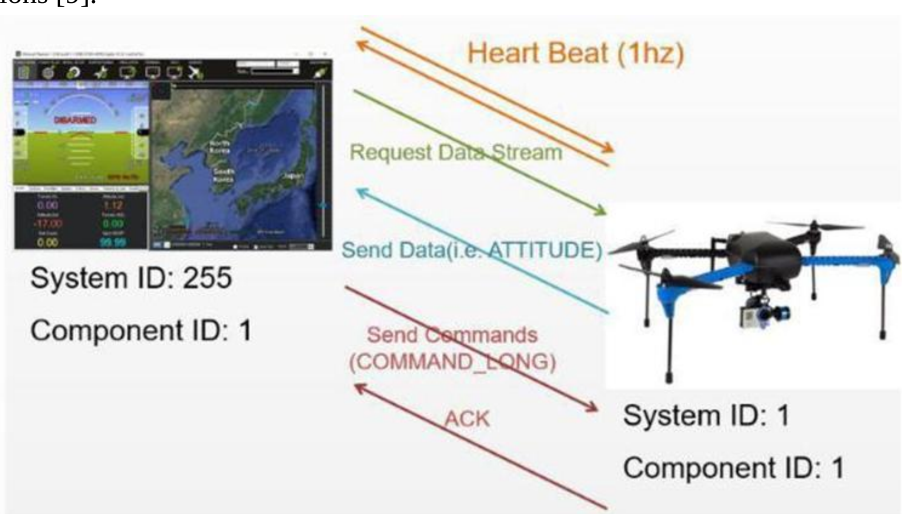

Figure 5: Mavlink connection between UAV and GCS

The protocol defines a large set of messages that can be sent over almost any serial connection and does not depend upon the underlying technology (Wi-Fi, telemetry, etc.). The messages are not guaranteed to be delivered which means ground stations or companion computers must often check the state of the vehicle to determine if a command has been executed [9].

2.3 Tracking Algorithm

After experimenting with different algorithms as mentioned above, we came up with a computationally efficient and robust algorithm, the discriminative correlation filter with Channel and Spatial Reliability. We bring the notions of channel and spatial reliability to DCF tracking, as well as a unique learning technique for their efficient and seamless integration in the filter update and tracking processes. The filter support is adjusted by the spatial reliability map to the part of the object that is suitable for tracking. This provides for a larger search area as well as better tracking of non-rectangular items. The channel-wise quality of the learnt filters is reflected in reliability scores, which are employed as feature weighting coefficients in localization. We initialize our video stream and begin looping over each frame of the video. We read a frame and resize the image size to improve our processing time. Then read the height and width of the resized frame. Next, we need to locate the object we want to track and draw a bounding box around it, to do this we call the select ROI function, which opens up a GUI to select our bounding box and assign a random color to our bounding box [1].

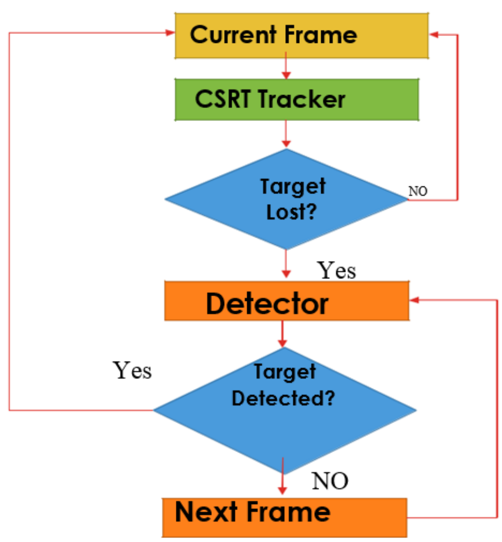

Figure 6: Steps in tracking algorithm

Now we will store the coordinates of the bounding box in a variable and initialize our CSRT tracker since it produced the best results. Our tracker is initialized using the frame and the location of the bounding box where our object to be tracked is as inputs. The tracker then passes this information to loop on every frame.

With this our tracker is initialized and ready to track, now we will use the update method to locate the position of our object in every successive frame and draw the bounding box around the updated position. Also determines its position in the frame and determines if the object is moving out of our central region of interest (Fig. 8) then generates commands of left, right, top, and bottom, etc. These commands then send to Pixhawk using mavlink, mavproxy protocols to generate UAV motion [8]. When we initialize our algorithm, CSRT tracking algorithm is called CSRT.

Also knows as Discriminative Correlation Filter with Channel and Spatial Reliability (DCF-CSR) [6] uses a spatial reliability map to adjust the region of focus of our correlation filter, this allows a better localization of the selected region and results in more accurate tracking especially in the case of non-rectangular regions of objects. The second thing this tracker does is in the channel reliability, which is estimated from the properties of the constrained least squares. The result of this is used to weigh the filtered response of each channel in localization.

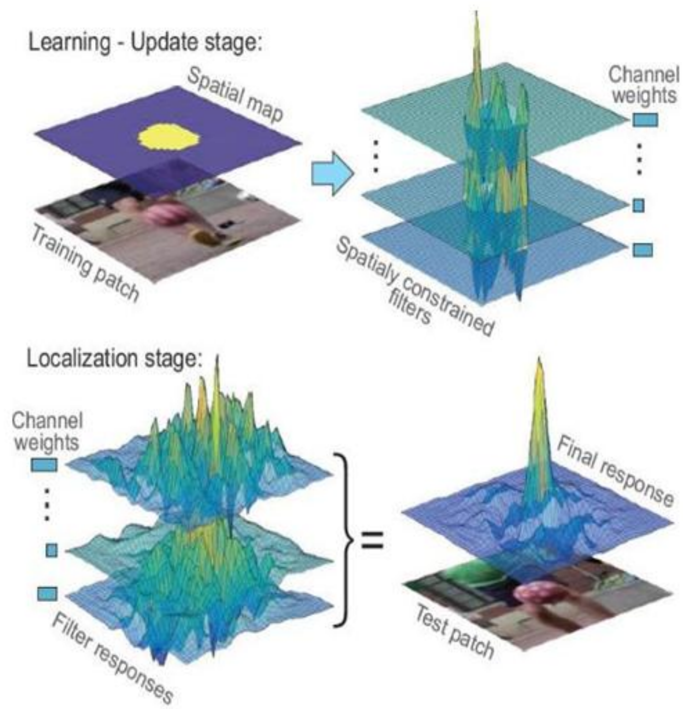

Figure 7: Overview of proposed CSR-DCF approach Steps in Tracking Algorithm

From Fig 7. The spatial reliability map is used to restrict the correlation filter to regions most suitable for tracking (top) resulting in an improved performance and search range for irregularly shaped objects. While the reduction of noise in the filtered response (bottom) is accomplishing using the channel reliability weights calculated during the optimization process.

2.4 Limitations of CSRT

Though CSRT works fine and gives good results then MOOSE and KCF but it has some limitations like slower FPS but this overcome by our powerful GPU processing computer [2], [10], [7]. Secondly, it loses track when the person gets out of the frame and it is hard to redetect object so to cater to this, we build a center region of interest, as described above, to keep our object in that region we move our drone as our objects gets out of that box. Comparative to detection algorithms, tracking algorithms it has errors like when an object is under a tunnel or object is covered by something for an extended period of time so to counter this usually a detection algorithm is trained on specified objects, so they have more knowledge classification of objects but currently, we are reinitiating our tracker to solve this.

3. Results

From experimenting algorithm on a dataset of UAV videos and live video captured from our drone. We concluded that our algorithm gives good tracking accuracy and fps rate unless our target object gets out of frame.

As we described in the limitations section that CSRT loses track when objects get out of the frame it loses track, so we solved this by creating a region of interest (green color box). The green box in the above images shows our center region of interest where we want our object to be and the bounding box around the object shows the target object which we are tracking. We can see when an object is moving out of the region of interest (green color box) a respective command of flight motion is generating in the direction of object motion.

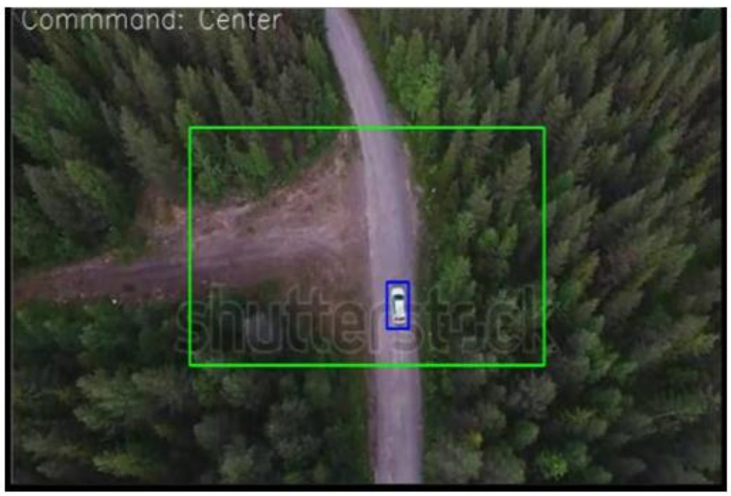

Figure 8: Result on dataset video

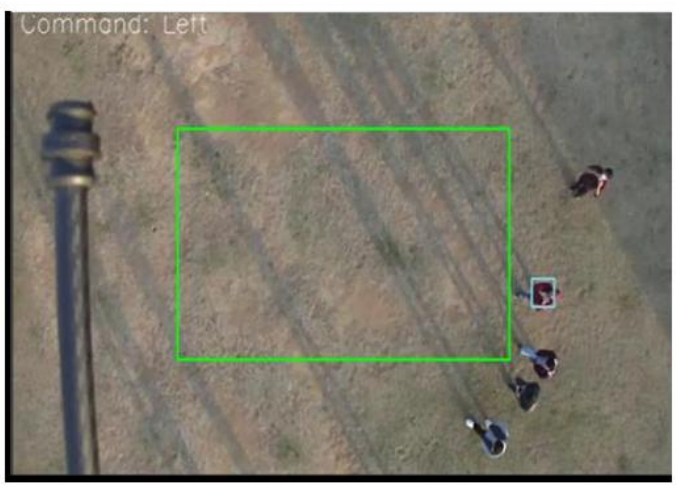

Figure 9: Result on live drone video

4. Conclusion

In this paper, robust object detection and tracking approach using a UAV mounted with a camera are presented. The proposed algorithm uses machine learning coupled with Channel and spatial reliability as its base. Furthermore, after the initial frame the tracker is bound to a region surrounding the object in the previous frame this allows limiting of the computational load while decreasing the number of false positives when dealing with multiple objects similar to our target in a frame and changes in illumination conditions and background. A dataset consisting of videos from the Internet and flight tests were processed to analyze the performance in terms of frame rate, detection, and tracking accuracy along with success rate. The target was detected, and the track was successfully maintained with adequate accuracy in various scenarios and terrains. Future works will be aimed at increasing the tracking accuracy while further optimizing the algorithm to improve frame rate and reducing the computational load even more.

References

[1] Hui Cheng, Lishan Lin, Zhuoqi Zheng, Yuwei Guan, and Zhongchang Liu, "An Autonomous Vision-Based Target Tracking System for Rotorcraft Unmanned Aerial Vehicles," IEEE/RSJ International Conference on Intelligent Robots and System, September 24–28, 2017.

[2] Wenbo Huang, KyoungYeon Ki, Yong Yang, Yoo-Sung Kim, "Automatic Shadow Removal by Illuminance in HSV Color Space", Computer Science and Information Technology, pp:70-75, 2015.

[3] Darshana Mistry, Asim Banerjee," Comparison of Feature Detection and Matching Approaches: SIFT and SURF", Global Research and Development Journal for Engineering , Volume 2 , Issue 4 , March 2017.

[4] Karami, Ebrahim & Prasad, Siva, Shehata, Mohamed, "Image Matching Using SIFT, SURF, BRIEF and ORB: Performance Comparison for Distorted Images" Newfoundland Electrical and Computer Engineering Conference, At St. John's, Canada, 2015.

[5] ZHAN Chaohui DUAN Xiaohui, XU Shuoyu, SONG Zheng LUO Min, " An Improved Moving Object Detection Algorithm Based on Frame Difference and Edge Detection", Fourth International Conference on Image and Graphics, 2007.

[6] Mantripragada, Kiran & Trigo, Flavio, Martins, Flavius & Fleury, A.T," Vehicle Tracking Using Feature Matching and Kalman Filtering" International Congress of Mechanical Engineering, At Ribeirao Preto, Brazil, Volume: 22, 2013.

[7] http://ardupilot.org/copter/docs/initial-setup.html

[8] https://docs.px4.io/en/flight_controller/pixhawk-2.html

[9] Ezio Malis, Manuel Vargas. "Deeper understanding of the homography decomposition for vision-based control." INRIA. pp.90, 2007.

[10] http://ardupilot.org/dev/docs/companion-computer-nvidia-tx1.html

[11] Samuele Salti, Andrea Cavallaro, Luigi Di Stefano, "Adaptive Appearance Modeling for Video Tracking: Survey and Evaluation", IEEE TRANSACTIONS ON IMAGE PROCESSING, May 17, 2012.

[12] Gonzalo R. Rodr´ıguez-Canosa , Stephen Thomas , Jaime del Cerro , Antonio Barrientos, Bruce MacDonald ," A Real-Time Method to Detect and Track Moving Objects (DATMO) from Unmanned Aerial Vehicles (UAVs) Using a Single Camera", Remote Sensing, April 20, 2012.

[13] Darshana Mistry, Asim Banerjee," Comparison of Feature Detection and Matching Approaches: SIFT and SURF", Global Research and Development Journal for Engineering, Volume 2 , Issue 4 , March 2017.