Detection of small object at a distance using flickering frequency and trajectory patterns

Sutthiphong Srigrarom[1*]

Received 20 November 2019, Revised December 2019, Accepted 11 January 2020

Abstract

In this proposal, we intend to research on the innovative and unconventional way to detect small object at a distance using flickering frequency and trajectory patterns (if possible). The object can be drones and can be under cluttered or obscured condition. The initial detecting part can be done by fusion of (1) flickering frequency pattern identification, (2) flight trajectory pattern (if possible) and (3) visual sensor with deep learning.

The subsequent tracking part is intended to be the ongoing-research 3D coordinated image-based dynamic localizations technique.

Keywords : flickering frequency pattern, flight trajectory pattern, visual sensor, deep learning

[1] National University of Singapore.

[*] Corresponding author, E-mail: spot.srigrarom@nus.edu.sg

1. Background

1.1 Drone detection

There are several techniques for detecting the drone, e.g. ambient RF signal, Radar, Acoustic Signal, Computer Vision and Sensor fusion. Figure 1 summarizes few good literature reviews on these techniques, especially X.F. Shi et al [1], Guvenc, et al [2] and Azari et al [3]. Here is the caption obtained from Guvenc et al [2]

Detection Technique	Advantages	Disadvantages
Ambient RF signals (e.g., [4, 6])	Low-cost RF sensors (e.g., SDRs), works in NLOS, long detection range. May allow deuthentication attacks for taking control of drone by mimicking a remote controller or spoofing GPS signal.	Need prior training to identify/classify different drones. Fails for fully/partially autonomous drone flights due to no/limited signal radiation from a drone/controller.
Radar (e.g., [7–10]	Low-cost FMCW radars, does not get affected from fog/clouds/dust as opposed to vision based techniques, can work in NLOS (more sophisticated). Higher (mmWave) frequencies allow capturing micro-Doppler/range accurately at the cost of higher path loss. Does not require active transmission from the drone.	Small RCS of drone makes identification/classification difficult. Further research needed for accurate drone detection/classification and machine learning techniques, considering different radar/drone geometries and different drone types which all affect micro-Doppler signatures. Higher path loss at mmWave bands limits drone detection range.
Acoustic signals (e.g., [11–13])	Low cost for simple microphones (cost depends on the quality of microphones). Can work in NLOS as long as the drone is audible.	Need to develop database of acoustic signature for different drones. Knowledge of current wind conditions and background noise is needed. May operate poorly under high ambient noise such as in urban environments.
Computer vision (e.g., [9, 14, 15])	Low cost for basic optical sensors. Pervasive availability of cameras even at most commercial drones that can be used as sensors.	Higher cost for thermal, laser-based, and wide FOV cameras. Requires LOS. Level of visibility impacted by fog, clouds, dust.
Sensor fusion (e.g., [9, 15])	Can combine advantages of multiple different techniques for wider application scenario, high detection accuracy, and long-distance operation.	Higher cost and processing complexity. Need effective sensor fusion algorithms.

Fig.1 Summary of literature review of drone detection

A team of Zhejiang university suggests to use sensor fusion (X.F. Shi et al [1]), which shows promising solution. Yet, the system is still effective only with the moderate to large drone sizes. All of these literatures conform conclusion that there is a gap for detecting small and fast drone, as the objects are either too small to be detected, need extensive training or high cost.

The work on background subtraction has been done in other application, e.g. traffic (Shahare et al [4]). To the team knowledge, there is no published work done on implementing or applying similar technique to drone for outdoor applications to date. Therefore, we can implement the similar system too.

Lastly, the flight trajectory tracking work was suggested by V-P Thai et al [5]. Their work is to classify the fixed wing from rotary wing aircraft. We intend to implement similar systems for more detailed classification of several type of flying objects.

1.2 Drone tracking

The visual motion or trajectory tracking of an object can be achieved through obtaining the relative motion of the object and the camera in a sequence of images. Optical flow tracking or motion-energy estimation are often considered as one category of visual trajectory tracking technique. The way these methods work is by extracting the velocity field and calculating the temporal derivatives of sequence images.

Trajectory tracking based on these methods can be fast, but its downside is that it is subjected to noise resulting in inaccurate values and it is unusable if pixels produced by the camera motion are more than a few.

Feature-based technique is also considered as another category for trajectory tracking. The way that feature-based technique works is by recognizing a single unique identifiable object or multiple identifiable structures across the scene in sequence images. This technique has its own pitfall as well. Tracking using single object identification are first detected and then estimating the camera position based on triangulation. However, the single object must be placed at different but known locations in the scene or environment. This means that knowledge of the geometric model of the environment must first be known. While identifying multiple structures across the scene operate with the same concept as a single object, this approach is highly dependent on the working environment.

There is need to for 3D object detection and dynamic trajectory tracking. There are some prior works done in this area. Papanikolopoulos et al [6] proposed idea of using visual tracking of moving target by a camera mounted on a robot. The work done was 2 dimensional and more on accuracy of visual tracking but not on object trajectory. Kawanashi et al [7] did 3D object detection by multiple active cameras. However, their system based on fixed camera in a lab with the effective range about 3m and was focused on the dynamic searching and focusing on the tracked object. Talukder and Matthies [8] worked on real-time detection of moving objects from moving camera using dense stereo and visual odometry. Their work paves way to do dynamic tracking, however, they use depth (dense stereo) camera for ground view and get the vehicle motion in 2-D ground plane only. Migliore et al [9] applied a single camera for Simultaneous Localization And Mapping (SLAM) with Mobile Object Tracking in dynamic environments. Their work focused on SLAM in lab and fixed camera environment. Nevertheless, they demonstrated that the ordinary monocular camera and be used for dynamic object tracking. Rodrguez-Canosa et al [10] used single camera with optical flow of the track object by with Kalman filter. They estimated the camera motion by creating an artificial optical flow field by estimating the camera motion between two subsequent video frames.

Scaramuzza et al [11], [12], [13], [14] did extensive work on computer vision and visual odometry for tracking of objects in various indoor conditions. Their methods show the possibility of using one camera to track objects in 2D with accurate trajectory and able to obtain feature information. They introduced the visual-inertial odometry to get the estimation of the camera pose. They also suggested to do the collaborative monocular SLAM with multiple UAV but for localizations of individual MAVs. Zhou et al [15] did dynamic objects in the urban environments and focused on object segmentation using Convolutional Neural Network but not on trajectory localization. Recently, Jing li et al [16] suggested to combine the optical flow for real time tracking the cars on the road from the airborne video via deep learning and digital homography. The 2D tracking of car is done by optical flow method. Their work gives way to compensate the camera motion by means of digital imaging, i.e. examining image scene in pixel level to estimate the camera motion. However, their method assumes linear translation of the camera between images, so the rotational motion of the camera is not accounted for. The proposed photogrammetry of the known scene in this proposal can cover both this translation and rotational motion of the camera.

2. Suggestion

In recent years, Unmanned Aerial Vehicle (UAV) usage has increased substantially ranging from commercial applications to personal hobby; mainly for aerial photography.

Though there are limitations and regulations for flying UAVs near prohibited areas, it is still tedious to control these flying spaces. Recently, there was an increase in drone intrusions that resulted in several cases such as drone intrusion at Singapore Changi International Airport (Fig.2) and London Gatwick Airport which caused several flight delays. Hence, having a robust system to detect and track an intruder- UAV when it enters the prohibited area is quintessential.

3. For detection

The challenges of the small and fact moving drone or object are the followings:

• Small object: The size of the object in range depends on the distance from the camera to the intruder-UAV. There is a situation that the intruder-UAV appears to be very small at the edge of the camera.

• Moving object at a distance or under Cluttered Background: Even though there are UAVs standing by to detect the target object from different angles, the intruder UAV and interceptor is moving at a distance and may travel high speed against the cluttered or obscured background.

The initial detection of the small object from a distance is crucial. Currently, there are several advanced deep-learning based visual technique to identify the object. However, the object must occupy or have certain number of pixels so that the training algorithm could recognize

Fig.2 Drone intrusion at Singapore Changi International Airport

the feature and be identified as drone. Although, the object might be large, e.g. 1m or in size, if such object is located far away, most of the visual sensor fail to identify as the object may appear too small in the scene. For example, the DJI Matric M100 (5kg and 60cm approximately) appear as just a dot, if it flies just about 100m away. For the even smaller drone, e.g. Parrot Tello (30cm) and fly at 50m away and fly fast appear just a line in the scene.

The flying object of the drone however have changed its appearance while travelling, either by its motion, or flickering patterns due to its rotary wings (rotors). Here, we propose the innovative way to detect, by indirectly observing the change of the object's appearance in the frame by means of change of 'event'.

As all the objects have unique flickering characters, e.g. waving tree branches and leaves, there are certain frequency associated with. Therefore, one can do stochastic regression for classification of the flickering frequency of object. This requires the pre-training of the known the flickering frequency & patterns. Then, from the recorded frequency changes, one can record and do flickering frequency analysis in the frequency domain to compare, classify and even identify, whether such flickering object corresponds to rotors, waving tree leaves, wind, etc. This is, somewhat, analogous to voice recognition. Also, this technique requires not many pixels to be used (in principle, even just 1 pixel can do). The advantage over the conventional visual technique is obvious. Besides, nowadays, there are event cameras (or even many other monocular cameras) that has high frame rate, and can respond quickly, i.e. able to capture very high frequency. Therefore, the idea could be realized and tested.

In this proposal, we intend to use this new flickering frequency technique as the initial means of detection. One can scan the entire scene and look for 'hot spot' where there is motion or change in pattern, which can be as small as 1 pixel in size, corresponding to small object at a distance away. Once the suspicious 'hot spot' is detected, one can still utilize the conventional camera to zoom (focus) in to get closer look at the object and do the usual deep-learning based visual technique to accurately identify the object whether it is a drone or something else.

Besides, as a small object might still be difficult to be identified, we can optionally check indirectly with another identifier -flight trajectory, if that is possible, e.g. when the object moves across the scene. Similarly, one can do training of the flight trajectories (paths) of various objects (e.g. fixed wing aircraft, multirotor UAVs, birds, etc.) to obtain the signatures, and do another stochastic regression for classification of the flight trajectory of object, to get better accuracy of the tracked object. Note that, as we investigate the overall flight trajectory, the limitation on fast motion of the object is not the issue, as compared to the normal pan-tilt-zoom observing camera.

In short, for detection, we aim at research on fusing 2 or more techniques (flickering frequency/change of scene, deep-learning based visual object detection and identification, and perhaps flight trajectory) to get better detection of the small and fast-moving drone. Our target is to be able to detect and identify the drone about 1m or less in size at more than 100m or more away. This is the challenge that current vision-based technology alone is still hard to achieve.

4. For tracking (Subsequent work)

After initial detection as just described, we can apply image (vision)-based object detector with deep learning to further identify, detect and track the targeted drone. We anticipate starting with YOLOv3, due to its fast inference time, accuracy and ability to detect small objects. As YOLOv3 with pretrained coco weights can detect 80 different classes, we can train YOLOv3 with inclusion of other architecture to specifically detect small UAV. Subsequently, we shall include the current state-of-the-art deep learning algorithm to improve performance and accuracy (shorter detection time, and better confidence scores).

Afterward, we follow up the detection by tracking the target with translation using a KCF via local search technique. We try both CNN features extraction and adaptive background subtraction method to analyse the robustness of small and high-speed object. Later, when the drifting appears or there are the changes in scale, we use a deep refinement network to refine the object boundary which results in the precise location of the objects in 3D space. In a single network evaluation, DRN can estimate the changes in scale even the object is small in the scene as well as compensate the drifting of the tracker by refining the object region estimated by the correlation filters.

The targeted tracked drone (intruder UAV) state estimator to be designed using the Unscented Kalman Filter (UKF). UKF allows estimation of nonlinear motion. Based on our previous works, using the information from the tracked drone and target object state estimator, we predicted its future trajectory using an exact dynamic model of tracked drone connected with a target object.

After object is detected and identified as a drone, the next step is to track, either for following & grasping or sensing & avoiding purposes. In this proposal, we propose to continue to use vision-based technique to track the object, by means of the ongoing-research coordinated imaged-based 3D localizations technique. In this technique, the object is tracked by the 2 or more cameras installed on either stationary or moving platform. The images from livestream videos to be processed via Deep-learning based (Deep Refinement Network)

It should be noted that the proposed visual tracking has advantage as it gives visual info in dimensions, trajectory, heading and even friendly or adversary thought its flying behaviour. The precise and robust tracking result in high-performance UAV collaborative planning. A common tracking approach which uses simple linear assumptions fail to track a fast-moving object in an unknown cluttered scene when the camera is shaking.

In addition, a strong wind can cause the unpredictable drift to the UAVs which have a high probability of lost in tracking. To solve the previous issues, we introduce a GPU-based visual tracking framework that integrates Kernelized Correlation Filters (KCF) with our customized Deep Refinement Network (DRN) for real-time onboard object tracking and recovery that could handle drifting, small and fast object tracking.

5. REFERENCES

[1] Xiufang Shi, Chaoqun Yang, Weige Xie, Chao Liang, Zhiguo Shi, and Jiming Che, Anti-Drone System with Multiple Surveillance Technologies: Architecture, Implementation, and Challenges, Digital Object Identifier: 10.1109/ MCOM.2018.1700430, IEEE paper no. 08337899, 2018,

[2] İsmail Güvenç, Farshad Koohifar, Simran Singh, Mihail L. Sichitiu, and David Matolak, Detection, Tracking, and Interdiction for Amateur Drones, IEEE paper no. 08337900, 2018

[3] Mohammad Mahdi Azari, Hazem Sallouha, Alessandro Chiumento, Sreeraj Rajendran, Evgenii Vinogradov, and Sofie Pollin, Key Technologies and System Trade-Offs for Detection and Localization of Amateur Drones, arXiv:1710.08478v1, 2017.

[4] Dipali Shahare 1 Ms. Ranjana Shende, A Survey on Moving Object Detection in Static and Dynamic Background for automated video analysiS, IJSRD - International Journal for Scientific Research & Development| Vol. 1, Issue 10, 2013

[5] Van-Phat Thai, Weixian Zhong, Thinh Pham, Sameer Alam and Vu Duong, Detecion, Tracking and Classification of Aircraft and Drones in Digital Towers using Machine Learning on Motion Patterns, IEEE paper no. 08735240, 2019

[6] N. P. Papanikolopoulos, P. K. Khosla, and T. Kanade, "Visual tracking of a moving target by a camera mounted on a robot: A combination of control and vision," IEEE transactions on robotics and automation, vol. 9, no. 1, pp. 14–35, 1993.

[7] T. Kawanishi, H. Murase, and S. Takagi, "Quick 3d object detection and localization by dynamic active search with multiple active cameras," in Object recognition supported by user interaction for service robots, vol. 2. IEEE, 2002, pp. 605–608.

[8] A. Talukder and L. Matthies, "Real-time detection of moving objects from moving vehicles using dense stereo and optical flow" in 2004 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)(IEEE Cat. No. 04CH37566), vol. 4. IEEE, 2004, pp. 3718–3725.

[9] D. Migliore, R. Rigamonti, D. Marzorati, M. Matteucci, and D. G. Sorrenti, "Use a single camera for simultaneous localization and mapping with mobile object tracking in dynamic environments," in ICRA Workshop on Safe navigation in open and dynamic environments: Application to autonomous vehicles, 2009, pp. 12–17.

[10] G. R. Rodr´ıguez-Canosa, S. Thomas, J. Del Cerro, A. Barrientos, and B. MacDonald, "A real-time method to detect and track moving objects (datmo) from unmanned aerial vehicles (uavs) using a single camera," Remote Sensing, vol. 4, no. 4, pp. 1090–1111, 2012.

[11] C. Forster, S. Lynen, L. Kneip, and D. Scaramuzza, "Collaborative monocular slam with multiple micro aerial vehicles," in 2013 IEEE/RSJ International Conference on Intelligent Robots and Systems. IEEE, 2013, pp. 3962–3970.

[12] C. Forster, M. Pizzoli, and D. Scaramuzza, "Svo: Fast semi-direct monocular visual odometry," in 2014 IEEE international conference on robotics and automation (ICRA). IEEE, 2014, pp. 15–22.

[13] D. Scaramuzza, M. C. Achtelik, L. Doitsidis, F. Friedrich, E. Kosmatopoulos, A. Martinelli, M. W. Achtelik, M. Chli, S. Chatzichristofis, L. Kneip et al., "Vision-controlled micro flying robots: from system design to autonomous navigation and mapping in gps-denied environments," IEEE Robotics & Automation Magazine, vol. 21, no. 3, pp. 26–40, 2014.

[14] C. Forster, Z. Zhang, M. Gassner, M. Werlberger, and D. Scaramuzza, "Svo: Semidirect visual odometry for monocular and multicamera systems," IEEE Transactions on Robotics, vol. 33, no. 2, pp. 249– 265, 2016.

[15] G. Zhou, B. Bescos, M. Dymczyk, M. Pfeiffer, J. Neira, and R. Siegwart, "Dynamic objects segmentation for visual localization in urban environments," arXiv preprint arXiv: 1807.02996, 2018.

[16] J. Li, S. Chen, F. Zhang, E. Li, T. Yang, and Z. Lu, "An adaptive framework for multi-vehicle ground speed estimation in airborne videos," Remote Sensing, vol. 11, no. 10, p. 1241, 2019.