Object Detection and Tracking on Low Resolution Aerial Images

Abstract: Object tracking from an aerial platform poses a number of unique challenges including the small number of pixels representing the objects, large camera motion, and low temporal resolution. Because of these unique reasons, low resolution aerial image analysis needs to be tackled differently than the traditional image analysis both in terms of the sensors, and computer vision algorithms. Recently, the Wide Area Motion Imagery (WAMI) sensor platform enjoyed increasing attention as it can provide reasonable spatial resolution single band imagery in addition to its large area coverage. Despite these advantages, there is still not enough sensory information and most WAMI systems struggle to persistently detect and track objects. Increasing the spatial resolution to record rich sensory information is a long-term goal to improve aerial image analysis. In the short-term, additional modalities, such as spectral data, can be key in identifying objects even in low spatial resolution and advances in sensor technology is starting to make limited hyperspectral data acquisition at video frame rates possible.

The sensor considered is called the Rochester Institute of Technology Multi-object Spectrometer, which is capable of collecting limited hyperspectral data at desired locations in addition to full-frame single band imagery similar to WAMI. By acquiring hyperspectral data quickly, tracking can be achieved at reasonable frame rates. More spectral samples can lead to a huge volume of data, so the relatively high cost of hyperspectral data acquisition and transmission need to be taken into account to design a realistic tracking system. By collecting and analyzing the extended (spectral) data only for the pixels of interest, we can address or avoid the unique challenges posed by aerial tracking. To accomplish this, we translate a traditional tracking-by-detection algorithm to aerial tracking domain and utilize convolutional features extracted from the hyperspectral data to boost tracking. Also, a non-linear bayes filter is integrated into the tracking pipeline to assist the proposed tracking-by-detection algorithm in handling the occlusions. The proposed system is evaluated on realistic, synthetic scenarios generated by the Digital Image and Remote Sensing Image Generation software. Additionally, another novel tracking-by-detection algorithm utilizing multiple correlation filters are proposed to perform embedded system compatible object tracking on consumer-driven drone data, called UAV123 dataset.

Similar to object tracking, object detection is another challenging task that requires reasonable visual cues for reliable performance. On the other hand, remote sensing and satellite images represent the objects with small number of pixels (0.1m - 3m ground sampling distance). Because of this reason, just like object tracking, object detection in aerial images needs to be handled differently than the object detection in traditional images. The Planet Labs, a satellite company, scan the whole earth twice daily, providing large data that can be used to design and train convolutional object detectors. One stage detectors (SSD, YOLO, RetinaNet) and two stage detectors (R-FCN, Faster-RCNN) dominate the object detection in MS-COCO dataset consisting of traditional images. However, their performance go down dramatically in low resolution satellite images. Therefore, the design of a novel network is required to better handle small objects. In this direction, a spatiotemporal analysis can be performed to better handle the small objects using their motion information. A Faster-RCNN like approach is then proposed to utilize the advantages of the spatiotemporal data to detect two objects, plane and ship.

Bio: Burak Uzkent is a Computer Vision Engineer at Planet Labs working on object detection in low resolution images. Previously, he held the Computer Vision Engineer position at Autel Robotics performing work on efficient object tracking at embedded systems. Prior to his industrial experiences, he pursued Ph.D. in the Chestar F. Carlson Center for Imaging Science at Rochester Institute of Technology. His Ph.D. thesis concentrated on object tracking on low resolution aerial images using an adaptive multi-model sensor. His research experiences lie in the fields of object detection and tracking in aerial and traditional images ,and image segmentation.

Homepage: https://buzkent86.github.io/