Occlusion Detection for Dynamic Adaptation

Zachary Mulhollan, Aneesh Rangnekar, Anthony Vodacek, and Matthew J. Hoffman

Rochester Institute of Technology, Rochester, NY, USA
zjm1400@rit.edu, aneesh.rangnekar@mail.rit.edu, vodacek@cis.rit.edu, mjhsma@rit.edu

Abstract. Occlusion is a common issue for object detection and tracking applications using a remote sensor platform, especially in complex urban environments where occlusions from buildings, bridges, and trees are frequent events. While occlusions are unavoidable, the events can be predicted to occur before the object of interest is obscured if there is prior knowledge of the observed environment. To aid in object detection and tracking tasks, we create an environment to map terrain and find obscured regions in the scene which helps with re-detecting objects once they are no longer obscured. We propose a dynamic data driven applications systems (DDDAS) framework for detecting occluded regions in an imaged scene by integrating streams of real data with a physics-based simulation model that updates based on the most recent images.

Keywords: occlusion detection · remote sensing · dynamic adaptation

1 Introduction

We utilize prior knowledge from open source resources to detect occlusions within a given scene, allowing us initialize a simulation of the scene before we collect real data samples and update the scene. We use OpenStreetMap (OSM) [9] for scene initialization to obtain geo-rectified terrain of a given real world location with dense OSM tags. These tags can be mapped into 3D modeling and rendering software (for example, Blender). The 3D environment is then used to synthetically image the scene with the Digital Imaging and Remote Sensing Image Generation (DIRSIG) model [2, 3, 13], where common land cover materials such as concrete, grass, and asphalt can be assigned to have hyperspectral reflectance spectra. DIRSIG is a versatile too. that it can produce simulations of many image modalities such as RGB, multispectral, and hyperspectral through the visible and infrared spectrum.

OSM provides a priori information about ground surface regions within the scene that cannot be imaged directly from a remote imaging platform's specified position. For example, an airborne camera viewing a road network from a non-nadir viewing angle may not have direct line of sight on the road if nearby buildings and vegetation are obscuring the road. If road material spectra (asphalt, concrete) are not detected in a region where OSM claims a road exists, we

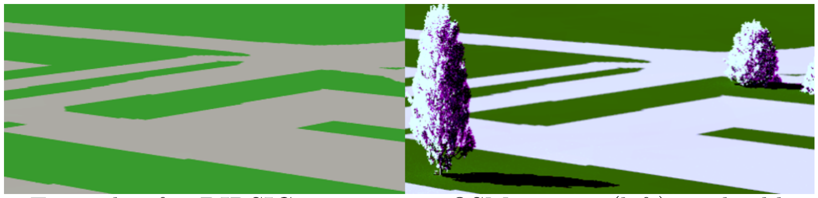

Fig. 1: Example of a DIRSIG scene using OSM terrain (left), and addition of trees (right) to the scene to occlude the ground terrain. Green represents ground vegetation and grey/white is paved roads.

infer the ground terrain in that region is occluded by an object that interrupts the airborne cameras line of sight, as shown in Fig. 1.

Our initial estimate of a simulated region using OSM contains information on a limited set of surface terrain materials such as asphalt and grass, which can then be confirmed or rejected with real image observations (Fig. 2). In a DDDAS sense the executing application is DIRSIG and new imagery are used to modify the DIRSIG inputs to modify the scene. We use the OSM information for constructing a scene and modify the scene with objects, like trees, to occlude the ground (Fig. 1). The proposed process will aid in object tracking systems from remote imagery, where objects moving in and out of occluded regions in a scene limits tracking performance. This paper considers the occlusion challenge in the task of detecting and tracking vehicles from a remote imaging perspective and uses scene simulation to overcome some of the challenge.

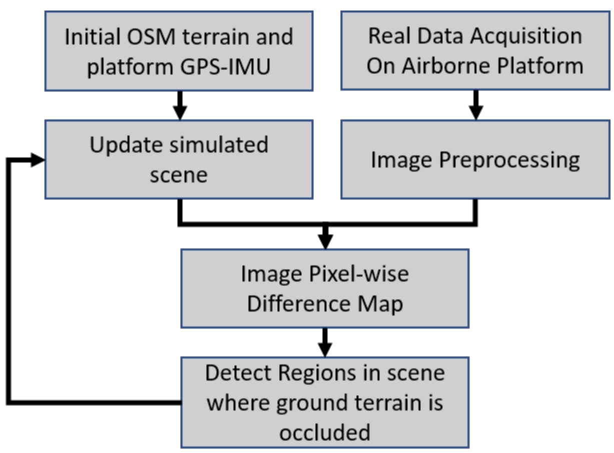

Fig. 2: Proposed framework for scene occlusion identification, with focus on roads and objects that may occlude the road network.

2 Related Work

Hang *et al.* used attention-aided CNN's (spectral and spatial sub-networks) for hypersepctral image classification of urban environments using the HyRANK dataset [5]. For moving object detection in aerial imagery, Palaniappan *et al.* utilizes background subtraction and depth mapping of tall structures that may occlude moving objects to reduce false positives due to parallax [10].

Since there are no publicly available video-rate hyperspectral datasets, simulated hyperspectral imagery is a viable method for constructing a hyperspectral dataset with a high framerate. There are handful of software approaches capable of creating hyperspectral imagery, such as: DIRSIG [2, 3], MCScene [12], CHIMES [15], and CameoSim [7]. We use DIRSIG to generate our simulated hyperspcetral imagery because it is a physics based renderer with an established history of publications, and it can accurately model radiation propagation through atmospheric modeling with MODTRAN [1].

Han *et al.* used DIRSIG to adjust atmospheric and environmental conditions for physics based data augmentation of simulated remote sensing imagery to train CNN's in vehicle detection [4]. Uzkent *et al.* modeled vehicle motion through a DIRSIG urban scene at various observation altitudes for object detection and tracking [14]. Kemker *et al.* used a DIRSIG desert scene to increase performance for semantic segmentation applications [6]. AeroRIT annotated all pixels in a hyperspectral aerial flight line over a college campus, initiating a baseline for use in hyperspectral semantic segmentation [11]. Mulhollan *et al.* collected the hyperspectral paint signature of over 450 vehicles using a calibrated drone mounted hyperspectral sensor, to aid in creating simulated hyperspectral imagery with a wide variety of vehicle reflectances [8].

3 On-the-fly Adaptations

Detecting and tracking a target vehicle with hyperspectral imaging through congested streets in an urban environment is a complex task. To aid in the task, we propose to utilize a-priori knowledge of the scene along with raw imagery to obtain additional geometric information from a 3-D perspective (**Solution 1:** Dynamic Metadata Integration). Metadata resources such as OpenStreetMap provide us with geo-rectified road layouts and land cover materials, which can assist in detecting occluded regions and provide probable locations for an occluded vehicle to reappear. The OpenStreetMap geographical land cover information initializes our scene, and the simulated model of the scene will update in regions of the scene where incoming real data are significantly different than the existing simulated model. Other sources of information such as the position and orientation of the imaging platform, position of the sun in the sky, and updated weather reports, all provide valuable information in predicting the expected spectral signature of the vehicle of interest.

3.1 Tackling Atmospheric Changes

A persistent bottleneck in object detection and tracking is the public availability of hypersepctral data. Hyperspectral cameras that can collect data at approximate video frame rates are rare and obtaining hyperspectral images from an airborne platform is costly and limited to flights in optimal weather conditions. Using physics based simulated imagery is a pragmatic method of acquiring a large hyperspectral video dataset of vehicles travelling through an urban environment. We use DIRSIG to generate our simulated hyperspectral imagery. This can alleviate data limitations and 1) provide pixel-wise ground truth data of all contents in the scene, 2) control for biases that often persist in real image data (such as weather condition and object orientation), and 3) provide a capability to create a scene with multiple imaging modalities. A simulated dataset also provides an automated emthod to tag environment based events such as the vehicle being occluded, shadowed, or contain glint, which we expect will make vehicle detection and tracking performance less dependent on atmospheric and environmental conditions.

Fig. 3 demonstrates the need to dynamically adjust the expected spectral signature of the vehicle as the scene changes over time. We observe a target vehicle in simulated hyperspectral imagery under two different weather conditions and five different airborne platform observation angles. Sunny afternoon observations of the light blue vehicle's spectral radiance are shown in orange, and same vehicle's spectral radiance observed in partly cloudy weather is shown in blue. The large difference in the signal amplitude demonstrates the dependence of target appearance on the illumination conditions (weather) and the angle at which the target is observed. Thus, it is important to update the expected target vehicle spectral radiance based on a-priori knowledge of the scene to improve performance of hyperspectral vehicle detection (**Solution 2:** Dynamic Signal Adaptation).

3.2 Dynamic Scene Reconstruction

Instead of an exhaustive training approach where a large hyperspectral dataset is collected of hundreds of vehicles from countless illumination conditions, observation angles, and occlusion events, we propose a DDDAS framework that utilizes physics based simulated hyperspectral imagery to predict how a target vehicle would appear to a real airborne imaging platform. We demonstrate that a physics based approach to hyperspectral vehicle detection can reliably locate a vehicle in complex urban environments, where illumination changes and occlusion events can often occur.

For occlusion detection, we first use OSM and the IMU-GPS positioning data of the aerial hyperspectral imaging platform to create a bare physics based simulated scene, where only on the ground materials such as grass, roads, walkways, and building footprints are geo-spatially placed based on OSM tagging. We simulate our image to look as if there were no vehicles or vertical occluding objects such as trees or buildings present in the scene. We ignore the land topography and use a planar surface to represent the ground for simplicity, but provisions are available to account for major changes in the topography.

In the real world, our airborne imaging platform collects a hyperspectral image of the scene, which may include any number of vehicles and occluding objects that are currently not populated in our physics based model. We update

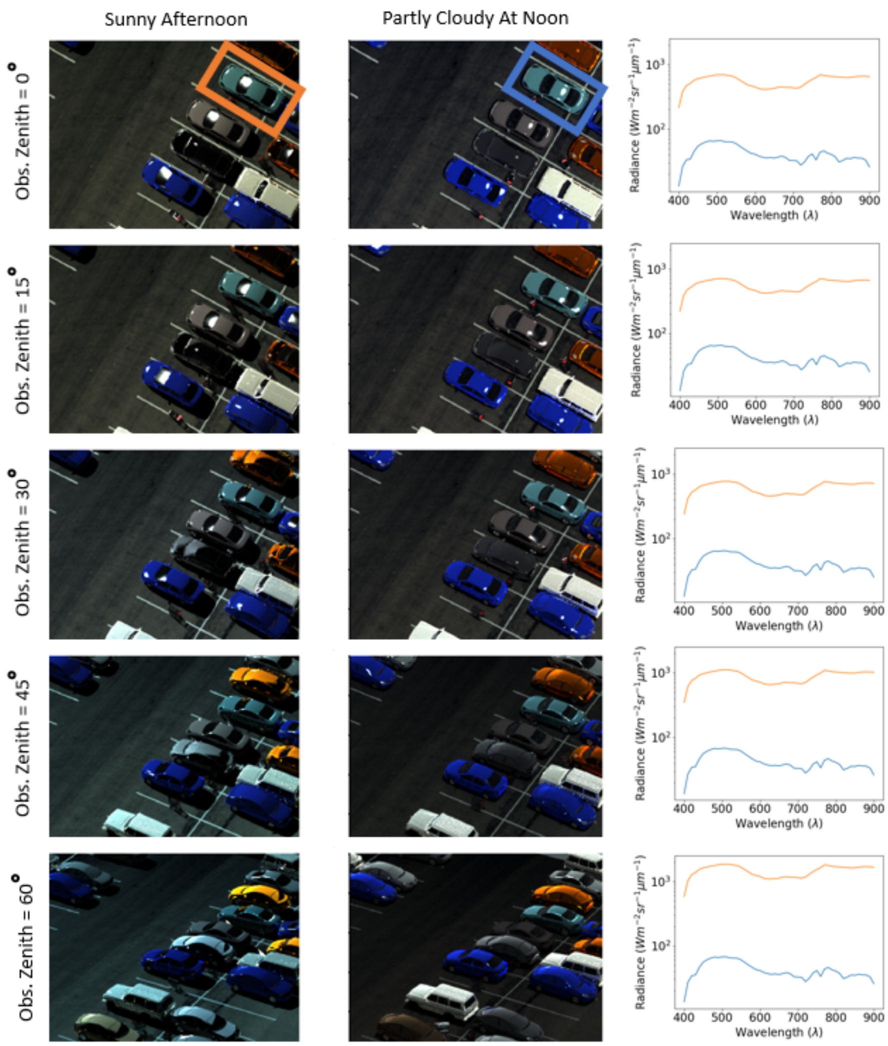

Fig. 3: Demonstration of a vehicle's spectral radiance dependence on atmospheric/weather conditions and image observation angle. It is also shown that expected glint on the vehicle can be modeled with accurate simulations of atmospheric conditions and observation angle.

the position and orientation of our simulated platform to best match with the latest position of the real image platform when the last frame was captured. We also process the data to geo-rectify the image and convert the pixel values from digital counts to physical units such as spectral radiance (**Solution 3:** Dynamic Scene Renderings).

To convert the data to physical units, we perform a lab calibration of the sensor to measure its spectral responsivity curve. The spectral responsivity curve paired with updated capture parameters on the airborne platform such as integration time, dwell time, and dark current provide enough information to calculate sensor reaching radiance

$$L_\lambda = (DC_i - DC_{dark}) \cdot \frac{E_\lambda}{DC} \cdot \frac{1}{\pi}, \tag{1}$$

Where L_λ is spectral radiance at the sensor, $\frac{E_\lambda}{DC}$ is the spectral irradiance per digital count obtained through a lab measurement of the sensor's responsivity curve, and DC_i and DC_{dark} are digital counts with incident light and dark current respectively.Spectral calibration of the sensor and converting image data to physical units such as spectral radiance allows us to compare our simulated physics based image with the imagery captured with a real hyperspectral sensor.

4 Results

To detect and track a vehicle using its spectral radiance in a scene containing occulsions, we use a DDDAS framework to predict occluded regions in the scene and update where and how we expect the target vehicle to look based on our knowledge of the scene. Predicting occluded regions in a scene is an iterative process as more data are collected. We provide a visualization of how the simulated model of the scene can update through new observations by using a simulation as an example. In Fig. 4 we show the ground truth change detection from the initial OSM landcover simulation alongside a supervised classifier spectral angle mapping image for change detection. The spectral classes used are asphalt and vegetation with the spectral data of these classes sourced from real hyperspectral imagery acquired from the same geometric location of the simulated scene.

Fig. 4: Simulated DIRSIG scene with trees (left) and an occlusion and shadow mask ground truth image (center) showing all occluded and shadowed pixels in the scene in white, with spectral angle classification used to detect change between OSM terrain and the simulated image.

To evaluate object detection and tracking performance in dynamic adapting environment that is full of occlusions, we construct a simulated dataset that contains labeled ground truth information such as the paint color, vehicle make and model, the location of the vehicle using bounding boxes, and also we tag if the vehicle is occluded or visible for each image frame. This simulated dataset is useful to reinforce the logic of occluded objects such as vehicles because it provides examples of vehicles that exist but are not currently visible due to occlusion. We can utilize this simulated dataset along with our existing knowledge of the scene (road networks and occluded regions) to learn where to look in the image to redetect an occluded object, and with a DDDAS feedback loop can guide the airborne platform to a new location in the sky where it has a higher probability of detecting the object unoccluded.

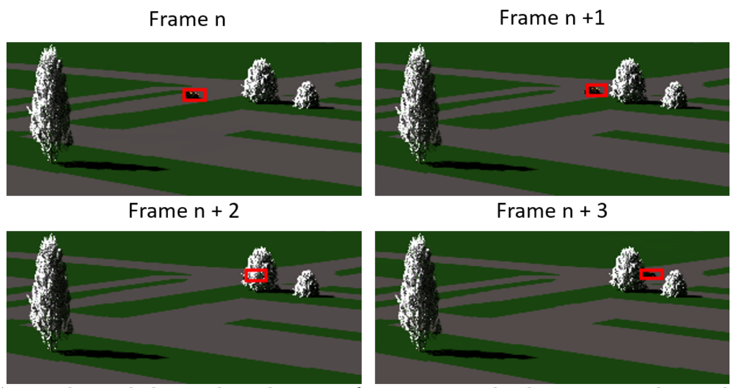

Fig. 5: Ground truth bounding boxes of a target vehicle moving through a simulated scene with occlusions caused by trees. This provides ground truth location of vehicles that are not directly observable due to line of sight obscurations.

5 Conclusion

We use a DDDAS approach to dynamically update a physics based hyperspectral simulated scene to the presence of occluded regions as new image information and metadata are provided. Detecting occluded regions in a scene aids object tracking and detection applications in complex urban environments, where moving objects vacillate between being obscured and visible. For hyperspectral detection of vehicles, we use simulated imagery to predict the expected signature of the vehicle's surface with atmospheric modeling and known geometric position of the imaging platform. We also construct a labeled simulated hyperspectral dataset with bounding boxes around each vehicle present in the scene, including ground truth location of vehicles that are occluded and undetectable in the raw imagery. This dataset will be used to train dynamically adapting detection algorithms to make vehicle detection and tracking applications more robust to occlusions, and in DDDAS framework can reposition an airborne sensor to a line of sight where the target vehicle is no longer occluded.

References

1. Berk, A., Conforti, P., Kennett, R., Perkins, T., Hawes, F., Bosch, J.: Modtran® 6: A major upgrade of the modtran® radiative transfer code. pp. 1–4 (06 2014). https://doi.org/10.1109/WHISPERS.2014.8077573

2. Goodenough, A.A., Brown, S.D.: Dirsig 5: core design and implementation. In: Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XVIII. vol. 8390, p. 83900H. International Society for Optics and Photonics (2012)
3. Goodenough, A.A., Brown, S.D.: Dirsig5: next-generation remote sensing data and image simulation framework. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing **10**(11), 4818–4833 (2017)
4. Han, S., Fafard, A., Kerekes, J., Gartley, M., Ientilucci, E., Savakis, A., Law, C., Parhan, J., Turek, M., Fieldhouse, K., et al.: Efficient generation of image chips for training deep learning algorithms. In: Automatic Target Recognition XXVII. vol. 10202, p. 1020203. International Society for Optics and Photonics (2017)
5. Hang, R., Li, Z., Liu, Q., Ghamisi, P., Bhattacharyya, S.S.: Hyperspectral image classification with attention aided cnns (2020)
6. Kemker, R., Salvaggio, C., Kanan, C.: Algorithms for semantic segmentation of multispectral remote sensing imagery using deep learning. ISPRS Journal of Photogrammetry and Remote Sensing **145**, 60–77 (2018)
7. Moorhead, I.R., Gilmore, M.A., Houlbrook, A.W., Oxford, D.E., Filbee, D.R., Stroud, C.A., Hutchings, G., Kirk, A.: Cameo-sim: a physics-based broadband scene simulation tool for assessment of camouflage, concealment, and deception methodologies. Optical Engineering **40** (2001)
8. Mulhollan, Z., Rangnekar, A., Bauch, T., Hoffman, M.J., Vodacek, A.: Calibrated vehicle paint signatures for simulating hyperspectral imagery. arXiv: Image and Video Processing (2020)
9. OpenStreetMap contributors: Planet dump retrieved from https://planet.osm.org . https://www.openstreetmap.org (2017)
10. Palaniappan, K., Poostchi, M., Aliakbarpour, H., Viguier, R., Fraser, J., Bunyak, F., Basharat, A., Suddarth, S., Blasch, E., Rao, R.M., Seetharaman, G.: Moving object detection for vehicle tracking in wide area motion imagery using 4d filtering. In: 2016 23rd International Conference on Pattern Recognition (ICPR). pp. 2830–2835 (2016)
11. Rangnekar, A., Mokashi, N., Ientilucci, E., Kanan, C., Hoffman, M.J.: Aerorit: A new scene for hyperspectral image analysis. arXiv preprint arXiv:1912.08178 (2019)
12. Richtsmeier, S.C., Berk, A., Bernstein, L.S., Adler-Golden, S.M.: A 3-dimensional radiative-transfer hyperspectral image simulator for algorithm validation (2001)
13. Schott, J.R., Brown, S.D., Raqueno, R.V., Gross, H.N., Robinson, G.: An advanced synthetic image generation model and its application to multi/hyperspectral algorithm development. Canadian Journal of Remote Sensing **25**(2), 99–111 (1999)
14. Uzkent, B., Rangnekar, A., Hoffman, M.J.: Tracking in aerial hyperspectral videos using deep kernelized correlation filters. IEEE Transactions on Geoscience and Remote Sensing **57**(1), 449–461 (2018)
15. Zahidi, U.A., Yuen, P.W., Piper, J., Godfree, P.S.: An end-to-end hyperspectral scene simulator with alternate adjacency effect models and its comparison with cameosim. Remote Sensing **12**(1), 74 (2020)