UNIVERSITY of DAYTON

VISION LAB

Center of Excellence for Computer Vision and Wide Area Surveillance Research

Nina Varney
varneyn1@udayton.edu

Yakov Diskin
diskiny1@udayton.edu

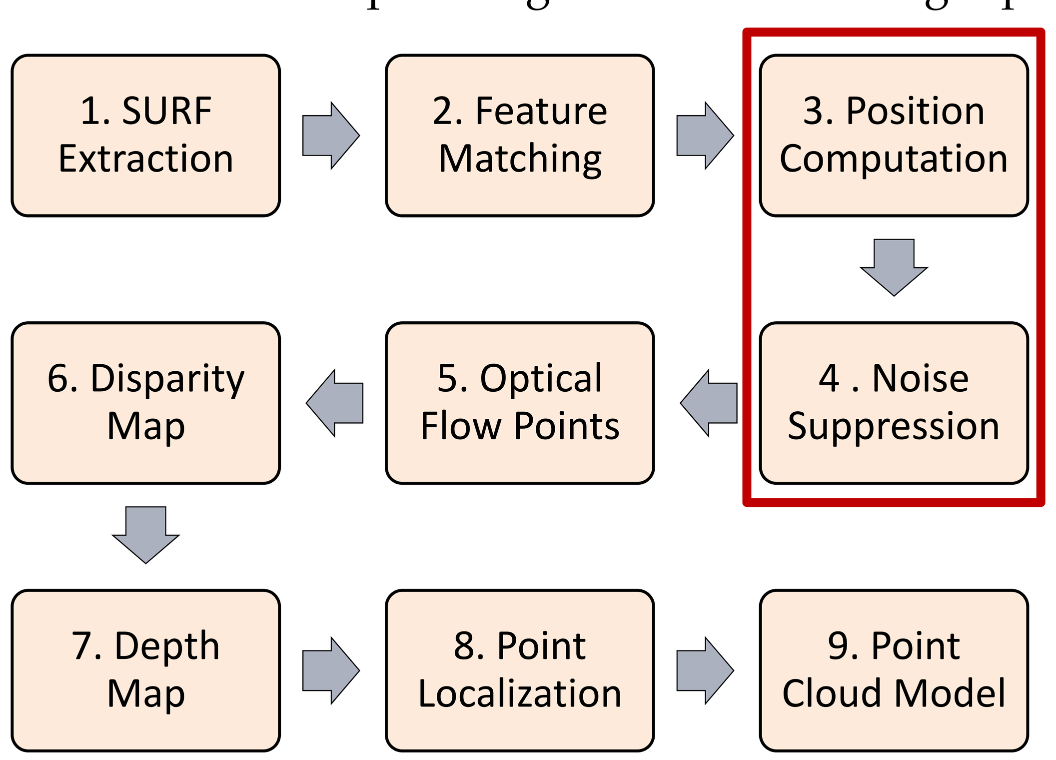
Classification of Vehicles using Monocular 3D Reconstruction

Introduction

State of the art 3D reconstruction techniques utilize frames from a video sequence to render a 3D model of the scene. Each point within the model has been tracked from frame to frame and triangulated into its (X,Y,Z) model position. We present an application for these structure from motion models that exploits our previous work in 3D object classification. In our experiments, we reconstruct a parking lot scene that contains several vehicles. The first step of our object classification algorithm is to segment each of the vehicles. Then, for each separate point cluster, our algorithm utilizes the volumetric and shape properties of the 3D object to label it with a vehicle type. The novelty of this classification approach allows us to tackle the noise challenges commonly associated with monocular 3D reconstructed models.

Structure from Motion Methodology

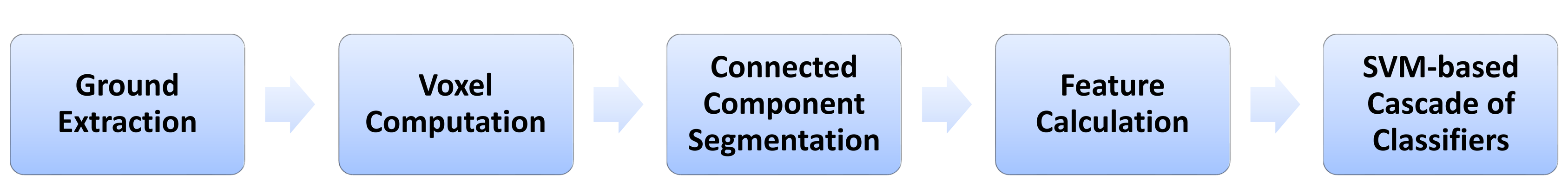
The nine step reconstruction process produces a Dense Point Cloud Representation (DPR) of a scene. Additional preprocessing steps can be added to enhance the input images and the resulting representation.



Visual Structure from Motion

Processing on un-calibrated imagery, Visual Structure from Motion (VSFM) computes the precise camera location and orientation associated with each frame. These computed parameters are essential when processing nonlinear camera motion.

3D Classification Methodology



The object space is considered and split into local subsections where a RANSAC based ground estimation is used to estimate the local ground plane. Once the ground points have been identified, the object points are split into a voxelized grid, where each element of the grid is represented by the density of points within that cell space. A three dimensional connected component approach is used to segment out individual objects.

Next, volume component analysis (VCA) is performed to extract reduced dimensionality features from each object, which represent the local direction variations. These VCA features will be used as the input to a cascade of support vector machines (SVM) to perform the final classification.

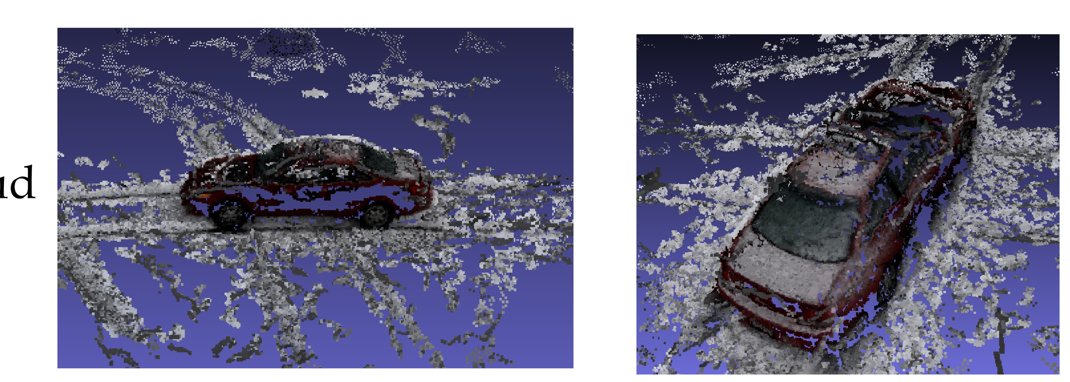
Data Collection

Many projects in the Vision Lab utilize aerial data, and the hexacopter drone is an effective way to obtain this aerial data. The TurboAce Cinewing 6 hexacopter carries a Canon 5D Mark III on board and has a flight time of up to 25 minutes. It is lightweight and easy to maneuver during flight making it an efficient way to collect aerial data.

Vehicle 3D Reconstruction

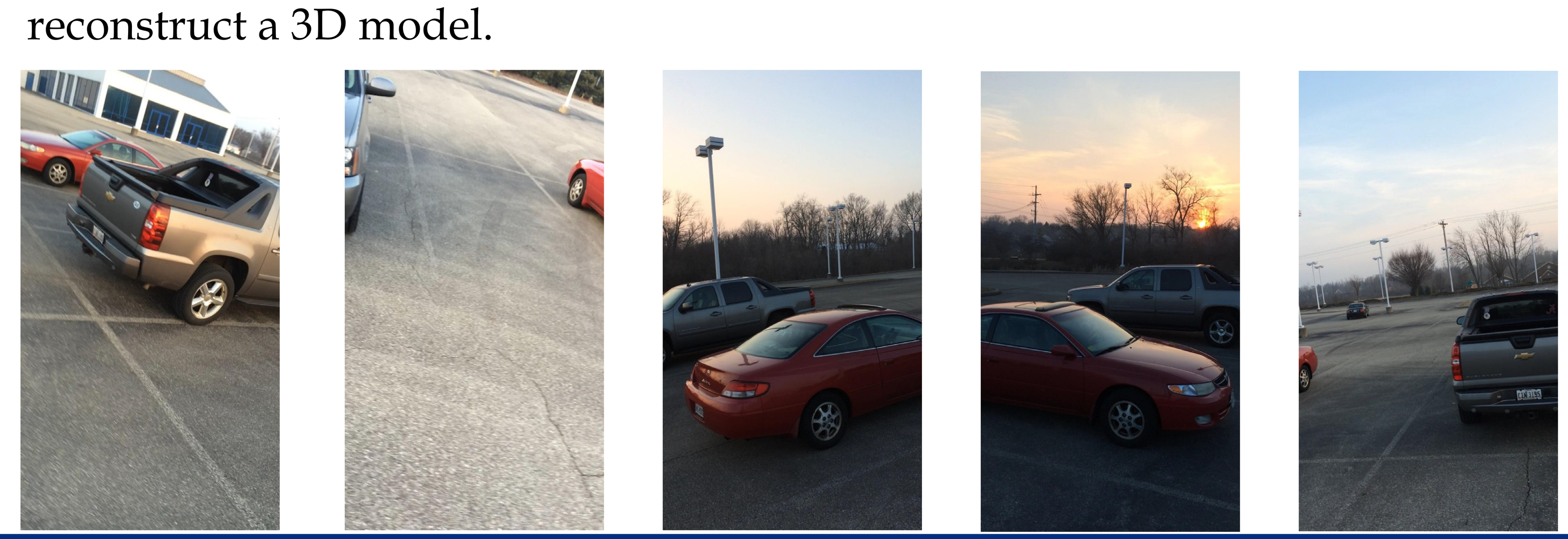
Input Imagery

3D Point-Cloud Model



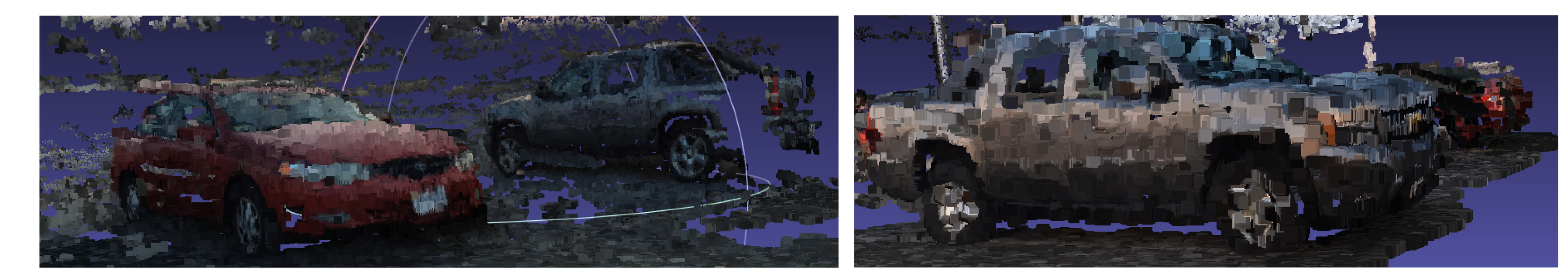
Dataset: Two Vehicles

The multi-vehicle data capture consisted of several cars parked in a lot. Aerial imagery of the scene was recorded and 40 frames of the scene were used to reconstruct a 3D model.



3D Vehicle Classification

After the 3D reconstructed point cloud is achieved we begin the classification. The classification strategy begins with a RANSAC ground estimation. The point cloud can be thought of as a scene with two main groups, ground points, which lie on the major ground plane, and object points, which vary in distance and density from the ground plane. Our goal in to use RANSAC plane estimation to find the points on this main ground plane. The results of the classification steps are shown below. We begin by observing the 3D point cloud models of the vehicles in the parking lot.



3D Object Segmentation

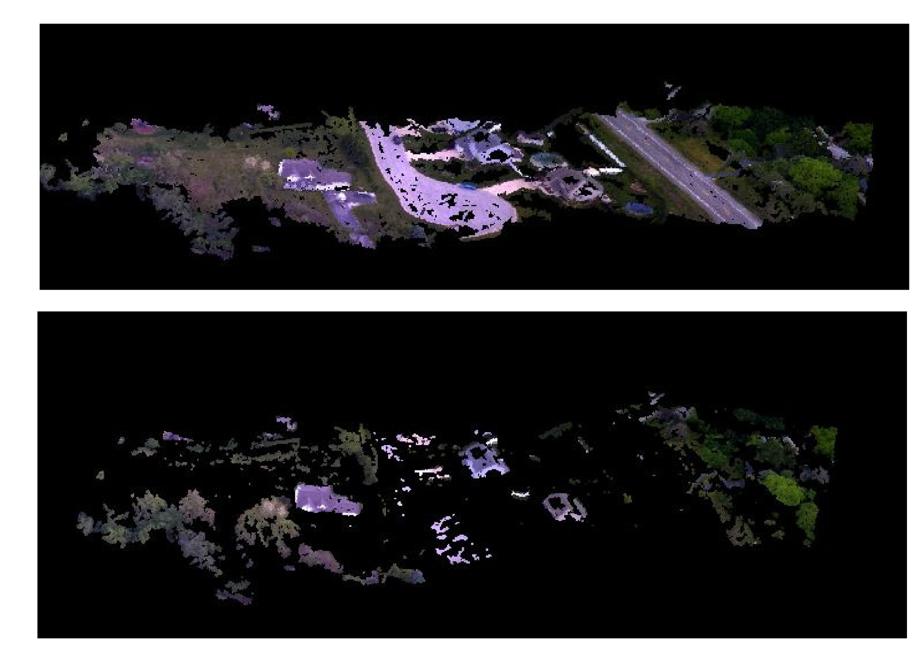
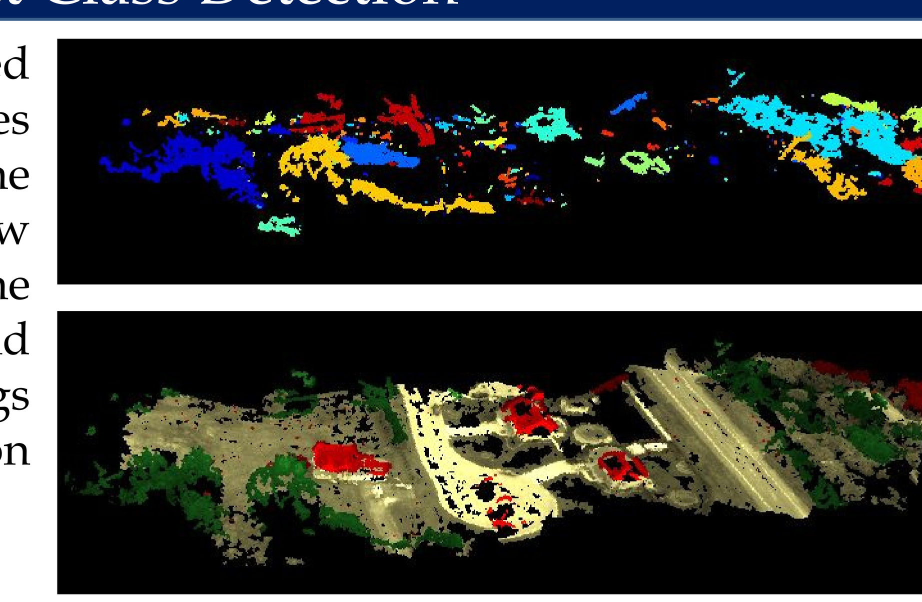


Figure to the left shows the original 3D reconstruction captured from an aerial platform, while figure to the left and below shows the segmented object points, after RANSAC ground estimation. The object in the scene have been separated into individual clusters. For each cluster we compute shape characteristics.

Object Class Detection

Each point cluster is classified based on volumetric properties and shape characteristics. The figure to the right and below shows the results of the classification, where the ground is shown in brown, the buildings are shown in red and vegetation is shown in green.



Continuing Work

- Distinguish between various models and makes of cars.
- Determine a proper model registration method to align the constant objects in the scene to improve the initial object detection process
- Apply noise suppression techniques to improve the accuracy of the classification.

UNIVERSITY of DAYTON

Vision Lab
Kettering Laboratories
300 College Park
Dayton, OH 45469-0232

UD VISION LAB
www.visionlab.udayton.edu

Dr. Vijayan Asari
Director
(937) 229-4504
VAsari1@udayton.edu