

Technologies for Safe & Efficient Transportation

THE NATIONAL USDOT UNIVERSITY
TRANSPORTATION CENTER FOR SAFETY

Carnegie Mellon University

UNIVERSITY of PENNSYLVANIA

Automated Video-Based Traffic Count Analysis

FINAL RESEARCH REPORT

CJ Taylor (PI), Yuting Yang, Univ. of Pennsylvania
Ryan Kennedy, Univ. of Pennsylvania, In collaboration:
Delaware Valley Regional Planning Commission

Contract No. DTRT12GUTG11

DISCLAIMER

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated under the sponsorship of the U.S. Department of Transportation's University Transportation Centers Program, in the interest of information exchange. The U.S. Government assumes no liability for the contents or use thereof.

Problem:

The goal of this effort has been to develop techniques that could be applied to the detection and tracking of vehicles in overhead footage of intersections. To that end we have developed and published techniques for vehicle tracking based on detecting and tracking point features. We showed how this approach could be combined with image based background subtraction to find and group vehicles. In subsequent work we showed how we could extract the 3D structure of a vehicle from a collection of feature tracks by exploiting the non-linearities associated with the pinhole projection model.

Approach:

We propose a method to extract traffic information from videos of road intersections using a vehicle detection, tracking, and 3D reconstruction system. The video is taken from a low-resolution camera fixed at a relatively low position, about 2 to 7m above ground. First, minimum eigenvalue feature points are detected and tracked with the Kanade-Lucas-Tomasi tracker to obtain feature point movements. These points are then segmented into groups corresponding to vehicles according to features such as position, movement and color. With the information of vehicle's grouping, the 3D position of each feature point can be calculated so that the vehicle's size, accurate position on road, time of entry and exit can be determined. The proposed method is robust to vehicles with varying movement and large perspective deformations.

Methodology:

There are several major difficulties that one can encounter in this kind of video analysis problem:

- Weather and lighting condition: The quality of the video depends on the weather and lighting conditions. Heavy rain will blur the scene making the vehicles harder to track. As will a lack of lighting during the night.
- Variety of movement: The video is captured at a road intersection, meaning that the vehicles in this video may stop and turn in manifold ways, which is far more unconstrained and unpredictable than the problem of analyzing highway traffic.
- Large perspective deformation and occlusion: The first problem is related to the pose of the camera. Unlike nadir views that would be provided by satellites or helicopters, the relatively low position of the camera will cause the vehicles to be larger when closer to the camera, and have distort the vehicles shape when turning. Also, occlusions often occur because of the low viewing angle of the camera. The other factor is due to the intrinsic parameters of the camera. The radial distortion and the position of the principle point will cause a simple homographic reprojection to fail to recover the "top view" of the video.

Our proposed method aims to solve the latter two difficulties. One approach to solving this type of problem is to use a feature-based tracking algorithm. Feature points are detected and tracked and then grouped according to various criteria, such as their position and current velocity. After grouping these feature points, traditional structure from motion techniques are used to recover the 3D information from the vehicles in order to acquire accurate information on vehicle's size and position.

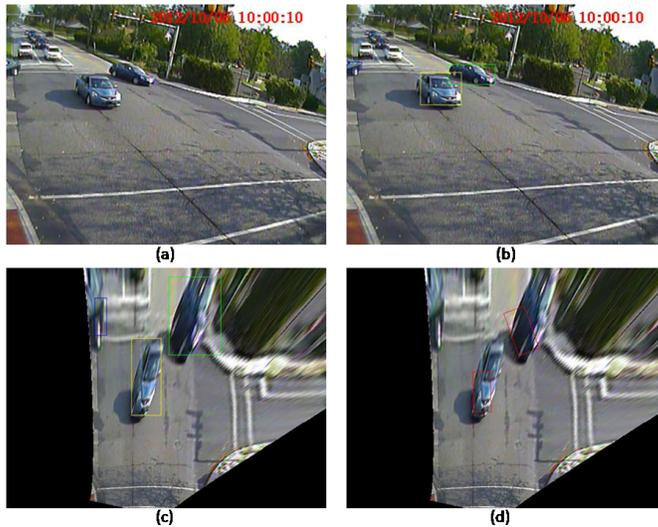


Figure 1: Sample frame (a) grouped in original perspective (b) and in "top view" perspective (c), and the size of the reconstructed vehicle represented by the red bounding box (d).

Our system contains three main steps: (a) determining the intrinsic parameters of the camera so that we can rewrap the image to an overhead view; (b) tracking the feature points and grouping them; and (c) recovering the 3D information for the grouped vehicle.

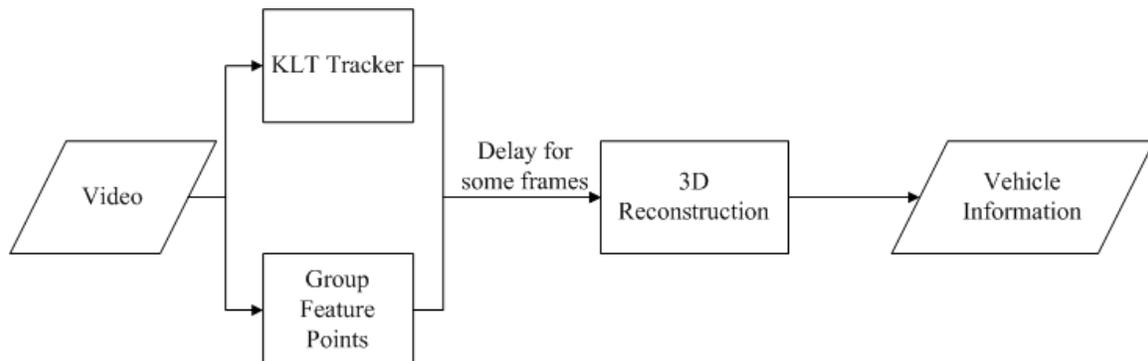


Figure 2: Block diagram of proposed approach.

Findings:

The algorithm was performed on six 10 seconds clip in the video from DVRPC. These clips covers different types of trajectories. To demonstrate the grouping and 3D reconstruction result, we sketched the outline of the road in a 3D plot as reference, then plot the 3D position of feature points and their bounding box on the road.

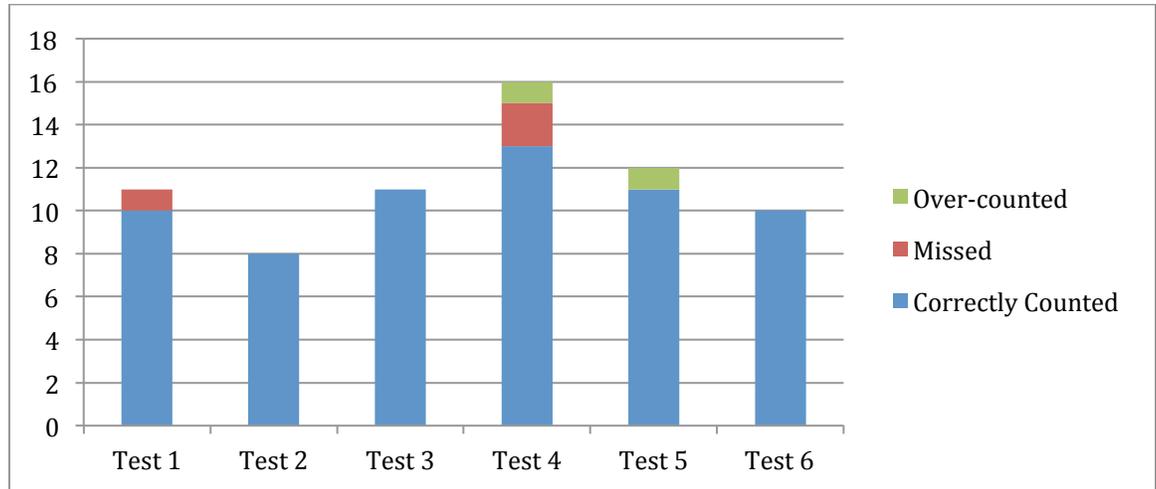


Figure 3: Performance of algorithm on different groups of vehicles in video tests.

Missed vehicles are mostly due to occlusion, and over-counted vehicles are mostly due to large vehicles like trucks and vehicles that are moving too fast to produce enough feature points. By tuning the parameters in the grouping step, we also computed a Receiver Operating Characteristic curve (ROC curve) between the true positive rate (recognized vehicles) and the false positive rate (over-counted vehicles).

Conclusions:

In this project, we developed a technique to combine the vehicle tracking and 3D reconstruction to determine traffic situations. This method is robust to perspective distortions and various vehicle movements. This system may be improved by feeding the 3D reconstruction back to the grouping stage to better predict occlusions between the vehicles which could be used to further improve the feature tracking procedure.

Recommendations:

There is still a need to further develop techniques for dealing with situations with significant traffic congestion where occlusion and segmentation issues become more challenging.