

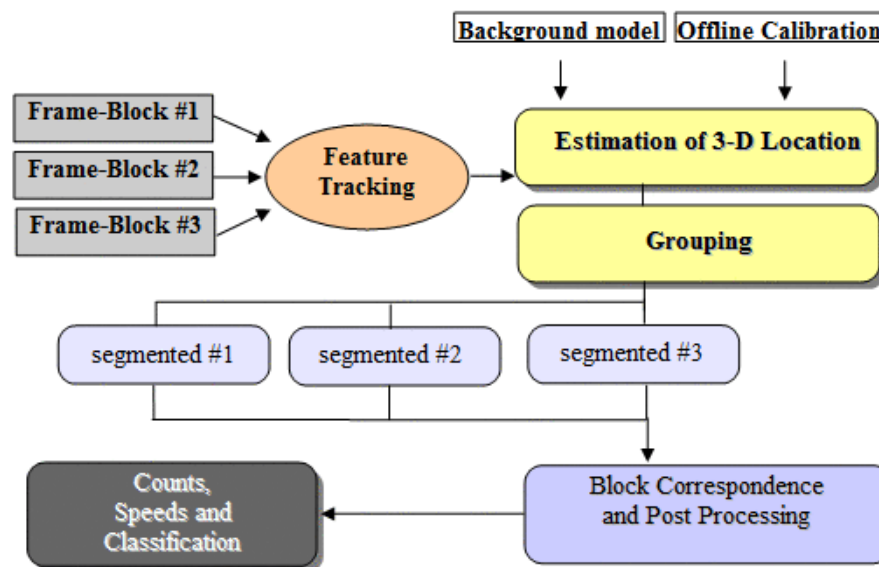
Vehicle Segmentation and Tracking in the Presence of Occlusions

Neeraj Kanhere (Electrical and Computer Engineering)
 Shrinivas Pundlik (Electrical and Computer Engineering)
 Dr. Stan Birchfield (Electrical and Computer Engineering)
 Dr. Wayne Sarasua (Civil Engineering)

Automatic traffic monitoring is becoming increasingly important as our need for highly reliable transportation systems grows. Among the many existing sensor technologies (e.g., inductive loops, hoses, infrared, ultrasonic sensors, magnetometers, and radar), video-based systems are emerging as an attractive alternative due to their ease of installation and operation, as well as their increased accuracy. Cameras are capable of capturing a rich description of the traffic parameters, including not only vehicle count and average speed but also parameters such as trajectories, queue length, and classification. Existing commercial solutions, however, have only begun to tap into this potential, with long setup times, expensive equipment, and impoverished extracted descriptions still common. Much of the research on developing algorithms for automatically extracting traffic parameters from cameras has focused upon the situation in which the camera is looking down on the road from a high vantage point, in which case the problem is greatly simplified because the problems of occlusion and temporal change in visual appearance are largely non-existent. It is not always feasible, however, to place the camera at a high vantage point. In this work we present an algorithm for automatically segmenting and tracking vehicles from a single camera that is placed relatively low to the ground, in which case significant occlusion between vehicles occurs.

Overview of the approach

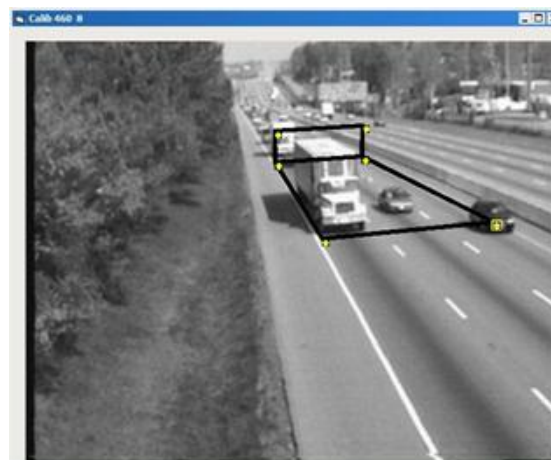
Our approach involves estimating the 3D coordinates of feature points using a two-layer homography, calibration parameters, and motion information::



Background model: A background image of the scene (with no moving objects present in it) is generated using a adaptive median filtering in the time domain. The number of frames required for generating a good background estimate varies with the amount of traffic in those frames. For the example shown above, 300 frames were used.

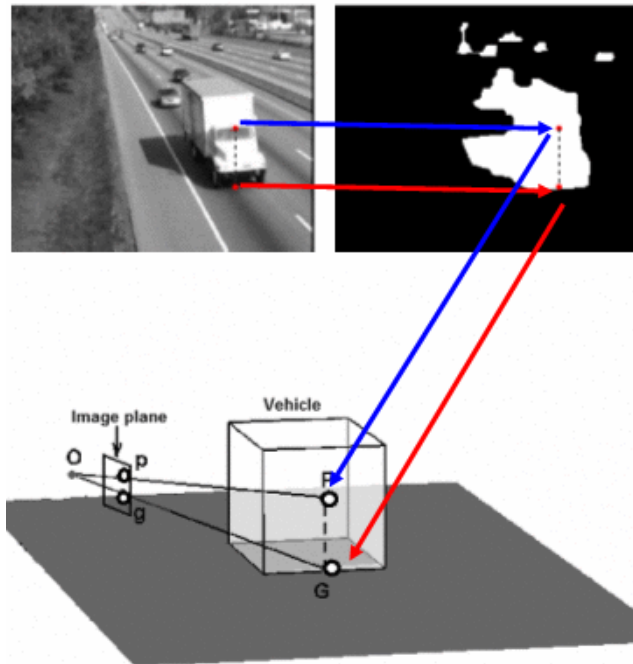


Offline calibration: To solve for the unknown projection parameters, we need at least six correspondences between world and image points. For accurate speed measurements the exact world coordinates are needed for the corresponding image points, but accurate segmentation requires only approximate calibration, which can be achieved using scene-specific features such as lane widths and dimensions of a known class of vehicles.



Feature tracking: Point features are detected and tracked in a *frame-block* (a set of consecutive frames) using the Kanade-Lucas-Tomasi (KLT) tracker.

Estimation of 3D locations: World coordinates for feature points are estimated using background subtraction and calibration parameters.

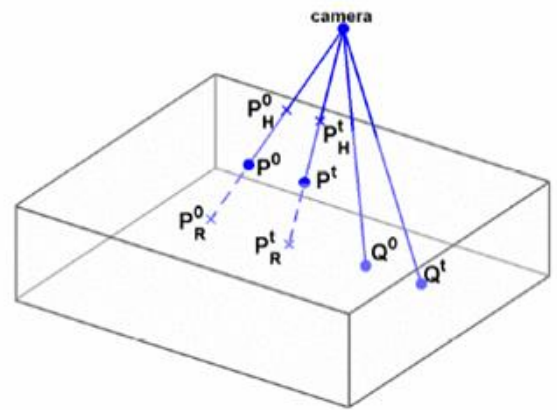


The subset of these features which are close to the ground plane are selected as *stable features* and used to estimate the world coordinates of rest of the features using motion information.

$$\mathbf{P}_k = \mathbf{P}_R + \frac{[\Delta_{P_H} - \Delta_{P_R}]^T [\Delta_{Q_k} - \Delta_{P_R}]}{[\Delta_{P_H} - \Delta_{P_R}]^T [\Delta_{P_H} - \Delta_{P_R}]} [\mathbf{P}_H - \mathbf{P}_R]$$

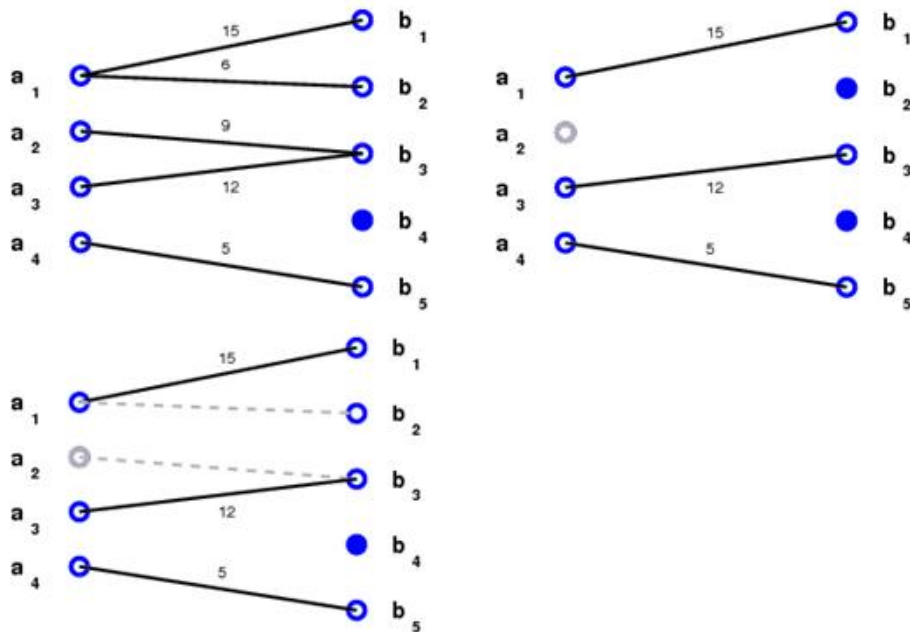
$$\mathbf{P} = \min_k \left\{ w_d \|\tilde{\mathbf{P}}_k - \tilde{\mathbf{Q}}_k\|_2 + w_e [\Delta_{P_k} - \Delta_{Q_k}]^T [\Delta_{P_k} - \Delta_{Q_k}] \right\}$$

- Coordinates of P are unknown
- Coordinates of Q are known
- R and H denote backprojections
- 0 : first frame of the block
- t : last frame of the block
- Δ : Translation of corresponding point



Grouping: An affinity matrix is computed based on the similarity between features. Shi and Malik's normalized-cut algorithm is used iteratively to group features.

Long-term tracking: Correspondence between feature groups that are detected and tracked in consecutive frame-blocks is established using the number of features that are tracked successfully between consecutive frame-blocks.



Results

The algorithm has been tested on several real sequences in which its ability to segment vehicles under severe occlusions is demonstrated:

- [Video 1](#) (MPEG2, 1 min 58 sec, 20 MB)
- [Video 2](#) (MPEG2, 38 sec, 6.5 MB)
- [Comparison](#) with a commercial detection-based system (MPEG2, 36 sec, 6 MB)
- [More results](#)

Some example image frames are shown below:



Data

The raw video sequences can be downloaded [here](#). All the AVI files are compressed using the xvid codec, which is included in the free [K-lite](#) codec pack. A program like VirtualDub (www.virtualdub.org) can be used to extract frames into

sequential jpeg/bmp files. Sequences were captured at 30 frames per second. In our experiments we used the first 10 minutes (18000 frames) for long sequences (L1, L2, etc.) and the first 30 seconds (900 frames) for short sequences (S1, S4, etc.). Unfortunately, sequences L6 and L7 are no longer available.

Publications

- N. K. Kanhere, S. T. Birchfield, W. A. Sarasua, and S. Khoeini. [Traffic Monitoring of Motorcycles During Special Events Using Video Detection](#). Transportation Research Record: Journal of the Transportation Research Board, 2010.
- N. K. Kanhere and S. T. Birchfield. [Real-Time Incremental Segmentation and Tracking of Vehicles at Low Camera Angles Using Stable Features](#). IEEE Transactions on Intelligent Transportation Systems, 9(1):148-160, March 2008.
- N. K. Kanhere, S. T. Birchfield, and W. A. Sarasua. [Vehicle Segmentation and Tracking in the Presence of Occlusions](#), TRB Annual Meeting Compendium of Papers, Transportation Research Board Annual Meeting, Washington, D.C., January 2006
- N. K. Kanhere, S. J. Pundlik, and S. T. Birchfield, [Vehicle Segmentation and Tracking from a Low-Angle Off-Axis Camera](#), Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), San Diego, California, June 2005
- N. K. Kanhere, [Vision-based detection, tracking, and classification of vehicles using stable features with automatic camera calibration](#), August 2008 (Ph.D. dissertation)
- N. K. Kanhere, [Vehicle segmentation and tracking from a low-angle off-axis camera](#), August 2005 (M.S. thesis)

This work was partially supported by a grant from the South Carolina State University Transportation Center.