

Probabilistic Obstacle Partitioning of Monocular Video for Autonomous Vehicles

Ryan W. Wolcott
rwolcott@umich.edu

Ryan M. Eustice
eustice@umich.edu

Ford Motor Company
Dearborn, MI
Perceptual Robotics Laboratory
University of Michigan
Ann Arbor, MI, USA

Localization is a key task for autonomous cars; systems such as the Google driverless car rely on precise and detailed maps for safe operation. Light detection and ranging (LIDAR) sensors are capable of providing rich information—including metric range and point appearance. Robust methods can use this data for vehicle localization by extracting the ground-plane for alignment to a prior map, as in [2].

Vision sensors as part of the localization pipeline can be a great enabler for autonomous platforms. Contrary to LIDAR methods, identifying the ground-plane from a camera image is a more challenging task. In our previous work [3], we considered localizing with a monocular camera by aligning the image to a prior map. As we demonstrated, this can be difficult as the ground-plane can be obscured by obstacles within view of the camera. In this work, we are interested in partitioning an image stream into obstacles and prior map, as shown in Fig. 1, so we can mask out obstacles during registration.

Similar to previous work [1, 4], we use a 1D-Markov random field (MRF) to model a horizontal image partition between obstacles and ground-plane, as in Fig. 1. However, rather than formulating our MRF potentials using image appearance alone (using learned [1] or hand-tuned features [4]), we instead consider the temporal stream of images and inferred parallax. We probabilistically evaluate optical flow against expected optical flow derived from known scene structure and camera egomotion, as in Fig. 2.

Our approach is evaluated on a challenging urban dataset with grayscale imagery, where lighting is non-uniform. We demonstrate our proposed algorithms by looking at errors with respect to hand-labeled groundtruth and present results showing improved image registration when obstacle masks are used.

Acknowledgements: This work was supported by a grant from Ford Motor Company via the Ford-UM Alliance under award N015392.

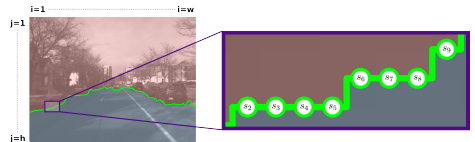


Figure 1: 1D-MRF to partition images into ground-plane and obstacles; each variable node in our MRF partitions an image column. Various unary potentials can be applied to each node; our work emphasizes a potential derived from optical flow.

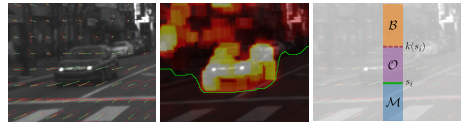


Figure 2: (Left) Optical flow vectors and expected flow vectors with uncertainties. (Middle) Optical flow likelihood and resulting partition. (Right) Optical flow potential implicitly considers segmenting image column into background (\mathcal{B}), obstacle (\mathcal{O}), and ground-map (\mathcal{M}).

- [1] Dan Levi, Noa Garnett, and Ethan Fetaya. Stixelnet: A deep convolutional network for obstacle detection and road segmentation. In *Proc. British Mach. Vis. Conf.*, pages 109.1–109.12, Swansea, United Kingdom, September 2015.
- [2] Jesse Levinson and Sebastian Thrun. Robust vehicle localization in urban environments using probabilistic maps. In *Proc. IEEE Int. Conf. Robot. and Automation*, pages 4372–4378, Anchorage, AK, May 2010.
- [3] Ryan W. Wolcott and Ryan M. Eustice. Visual localization within LIDAR maps for automated urban driving. In *Proc. IEEE/RSJ Int. Conf. Intell. Robots and Syst.*, pages 176–183, Chicago, IL, USA, September 2014.
- [4] Jian Yao, Srikumar Ramalingam, Yuichi Taguchi, Yohei Miki, and Raquel Urtasun. Estimating drivable collision-free space from monocular video. In *Proceedings of the IEEE Winter Conference on Applications of Computer Vision*, pages 420–427, Waikoloa Beach, HI, USA, January 2015.