Vehicle and Critical Event Detection in Night by Robust Stereo Analysis

AALBORG UNIVERSITY DENMARK

Mark Philip Philipsen & Morten Bornø Jensen

Advisors: Mohan M. Trivedi & Thomas B. Moeslund

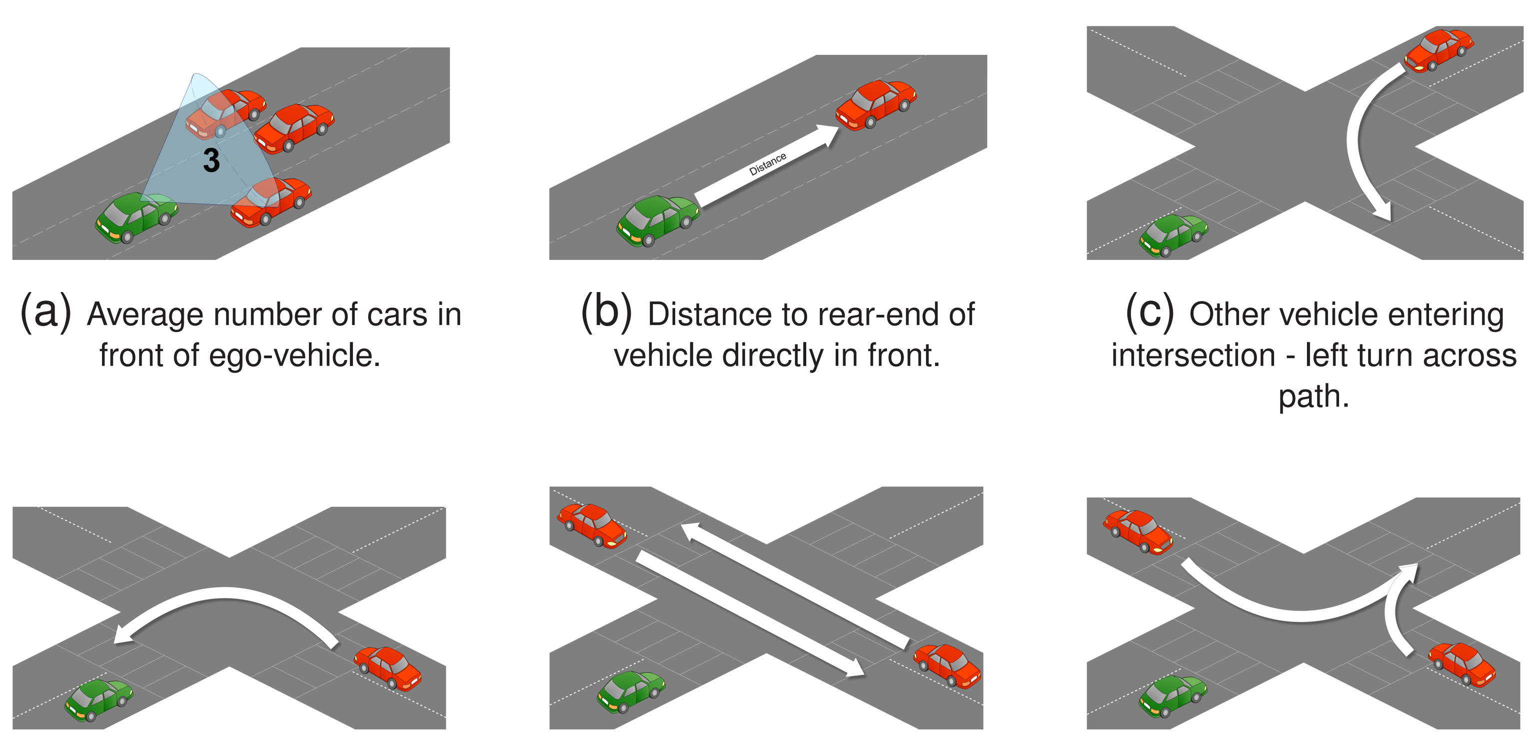
UC San Diego Jacobs School of Engineering

CVRR Computer Vision and Robotics Reserach Laboratory

LISA: Laboratory for Intelligent & Safe Automobiles

Introduction

We introduce a stereo-vision based system that enables automatic event detection based on detection and tracking of vehicles in scenarios that usually would be highly problematic for monocular detectors. The purpose is to provides insight in the patterns and behaviors of drivers during near-crashes and crashes. The proposed system will be limited to only handling a handful of events, which especially benefit from the extra dimension gained with stereo-vision. The events considered in this paper are illustrated in Figure 1. [1] Monocular systems usually have problems dealing with occlusions and are in some cases using classifiers based on appearance, which require a large amount of training data for dealing with detection of car from the many possible angles and light conditions.



(a) Average number of cars in front of ego-vehicle.

(b) Distance to rear-end of vehicle directly in front.

(c) Other vehicle entering intersection - left turn across path.

(d) Other vehicle entering intersection - turning onto opposite direction.

(e) Other vehicle entering intersection - straight across path.

(f) Other vehicle entering intersection - turning same direction.

Figure 1: Critical events that can be automatically detected by the proposed method. Green car is ego vehicle. Red cars are other vehicles.

Our contributions are:

- Using stereo-vision for automatic event detection in both day and nighttime, with focus on intersections (Figure 1c, 1d, 1e, 1f).
- Introducing a new event: Average number of vehicles in front of the ego vehicle. (Figure 1a).
- Introducing a new event: Average distance to vehicles directly in front of the ego vehicle. (Figure 1b).

Proposed System

A Bumblebee XB3 stereo camera is used to acquire stereo image pairs with an average rate of 11 FPS in an resolution of 1280x960. For generating the disparity map, the OpenCV's SGBM[3] implementation is used. Noise is removed using the The road surface is found by searching for the most significant line in the V-disparity[4] using RANSAC. Additionally, the line parameters are filtered using a Kalman filter to smooth out faulty road surface detections. Using the camera's focal length and baseline, along with the calculated disparity, the actual distance to objects in the camera's view can be found. The acquired point clouds are post processed using a band-bass filter to remove near and distant points, downsampled using a voxel grid, and removing outliers. Clusters are found by creating k-d tree, and for each point finding neighbors within a specified radius. The clusters center point is used for nearest neighbor tracking between frame and used to determine the distance from ego vehicle to detected vehicle.

For compensating for the subject vehicle's ego-motion while e.g. approaching an intersection, we utilize the *LIBVISO2: C++ Library for Visual Odometry 2* [2].

NDS events are detected by looking at the movement history of other vehicles with regard to the ego vehicle. For all detected vehicles, individual frame to frame movements are categorized to form a histogram of events for determining which NDS events have occurred.

An overview of the system is shown in Figure 2. The final output is an event report.

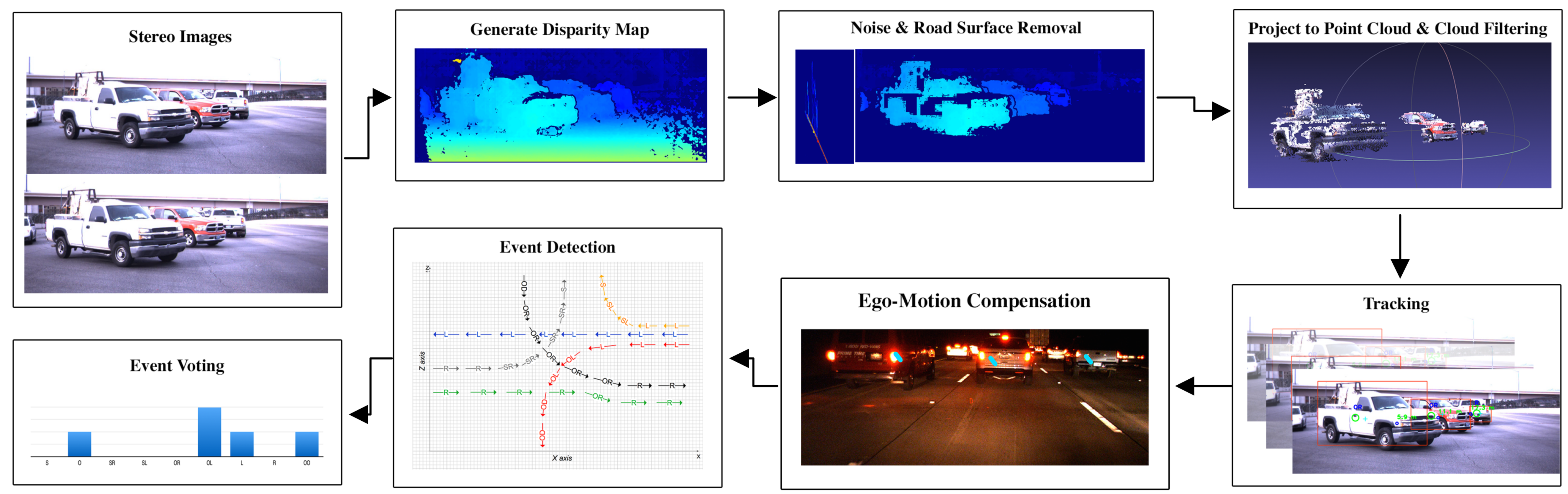


Figure 2: Overview of the stages used in the proposed system for automatic detection of critical events.

Results

The proposed system is evaluated on 216 day and 367 night time frames. In Table 1 the results are seen, where GT is the ground truth of events manually labeled and SO is the system output. The proposed system have an overall precision of 0.9375 and an overall recall of 0.71.

Table 1: Summary of event report analysis. The syntax for the results are [SO/GT]. P and R is abbreviations for precision and recall, respectively.

Drive Behavior Event	Daytime	Nighttime	P	R
Right - straight across path	34/38	8/15	1	0.71
Left - straight across path	14/21	2/4	1	0.58
Left turn across path	14/12	12/6	0.75	0.69
Turn onto opposite dir.	12/7	6/6	1	0.85
Avg. number of cars	1.67/1.74	1.6/1.3	NA	NA
Avg. distance to car	8.73 m	10.98 m	NA	NA

Concluding Remarks

The use of stereo-vision is considered beneficial, especially in scenarios with partly occluded cars, in such cases most monocular systems will fail. It is considered that providing distance information from the ego vehicle to surroundings provide useful information with regard to drive patterns before and during a crash or near crash. The system proved to work in both day and nighttime conditions with a limited drop in overall performance for the night scenarios.

Future work includes looking at additional events where stereo-vision can be utilized, evaluating on more data, and optimizing the system to work in real-time.

Acknowledgment

We would like to thank Ravi K. Satzoda and Andreas Møgelmose for their guidance.

References

[1] M. P. Philipsen, M. B. Jensen, R. K. Satzoda, M. M. Trivedi, A. Møgelmose, and T. B. Moeslund, "Night-time drive analysis using stereo-vision for data reduction in naturalistic driving studies," in *IEEE, Intelligent Vehicle Symposium*, 2015.

[2] A. Geiger, J. Ziegler, and C. Stiller, "Stereoscan: Dense 3d reconstruction in real-time," in *Intelligent Vehicles Symposium*, 2011.

[3] H. Hirschmuller, "Accurate and efficient stereo processing by semi-global matching and mutual information," in *IEEE CVPR 2005*.

[4] R. Labayrade, D. Aubert, and J.-P. Tarel, "Real time obstacle detection in stereovision on non flat road geometry through" v-disparity" representation," in *IEEE, Intelligent Vehicle Symposium, 2002.*