Vehicle and Critical Event Detection in Night by Robust Stereo Analysis

Mark P. Philipsen & Morten B. Jensen

Advisors: Mohan M. Trivedi & Thomas B. Moeslund

AALBORG UNIVERSITY DENMARK

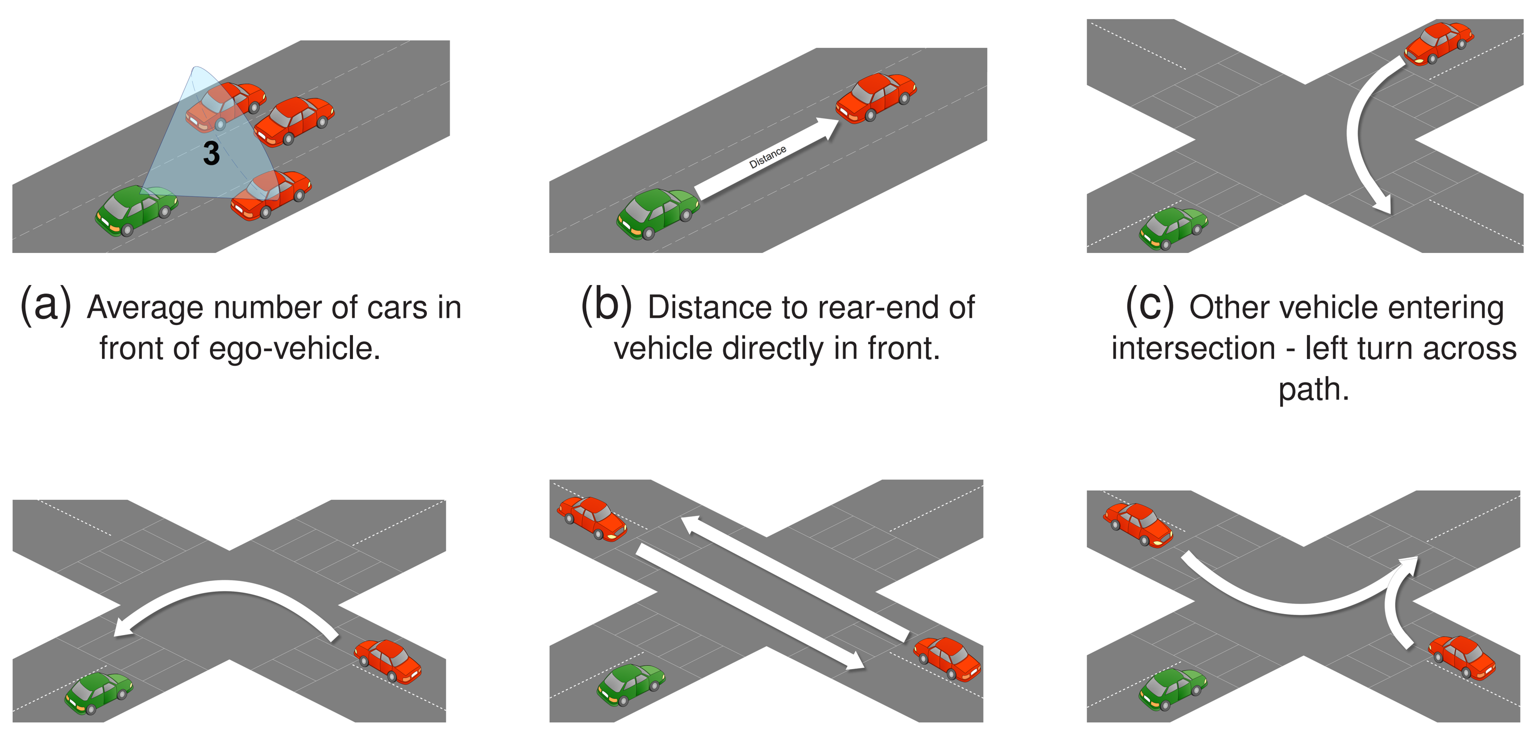
UC San Diego Jacobs School of Engineering

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 into 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. 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 vehicles from the many possible viewpoints 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.[1]

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 images** with an average rate of 15 FPS in an resolution of 1280x960. For generating the **disparity map**, the OpenCV's SGBM[2] implementation is used. **Noise** is removed using LR-RL consistency check, temporal information, and a monocular color check. The **road surface** is found by searching for the most significant line in the V-disparity[3] using RANSAC. Additionally, the line parameters are filtered using a Kalman filter to smooth out faulty road surface detections. Objects are projected to 3D using the camera's properties, which result in 3D **point cloud** representations of the segmented objects. The acquired point clouds are post processed using a band-bass filter to remove near and distant points, downsampled using a voxel grid, and outliers are removed. **Clusters** are found by creating a k-d tree, which organize points according to distance to neighbors.

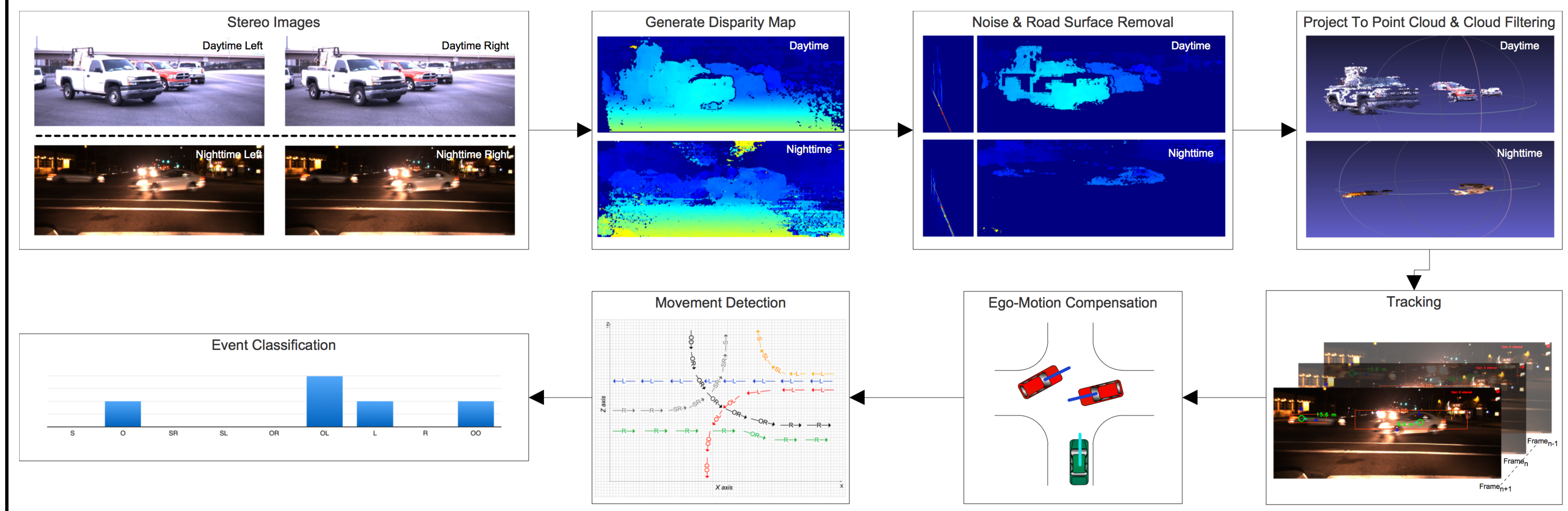


Figure 2: Critical event detection process flow.

The clusters' center points are used for nearest neighbor **tracking** between frames and for determining the distance from ego-vehicle to detected vehicles. For **ego-motion compensating** while approaching an intersection, we utilize the *LIBVISO2: C++ Library for Visual Odometry 2* [4]. Events are detected by looking at the **movement** history of other vehicles. For all detected vehicles, individual frame to frame movements are categorized to form a histogram of movements for **event classification**.

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

Results

The proposed system is evaluated on 4,992 day and 3,933 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 0.78 and recall of 0.72.

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

Drive Behavior Event	**Daytime**	**Nighttime**	**P**	**R**
Right - straight across path	35/32	5/19	0.95	0.63
Left - straight across path	45/34	11/33	0.87	0.67
Left turn across path	5/5	20/1	0.75	1
Turn onto opposite dir.	32/37	41/15	0.68	0.93
Short turn onto same dir.	7/5	9/5	0.63	1
Long turn onto same dir.	1/16	1/8	1	0.09
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. Distance information from the ego vehicle to the surroundings provide useful information with regard to drive patterns before and during a crashes.

- Introduction of a novel stereo based critical event analysis approach.
- Experimental analysis shows very promising detection, trajectory and event classification rates.
- Ongoing research involves extensive experimental validation and a day and night time critical event detection module for public use.

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

[3] 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.*

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

Acknowledgment

Thanks to Ravi K. Satzoda and Andreas Møgelmose for guidance.