Night-Time Drive Analysis using Stereo-Vision for Data Reduction in Naturalistic Driving Studies

AALBORG UNIVERSITY DENMARK

Mark Philip Philipsen & Morten Bornø Jensen

Advisor: Professor Mohan M. Trivedi

UC San Diego Jacobs School of Engineering

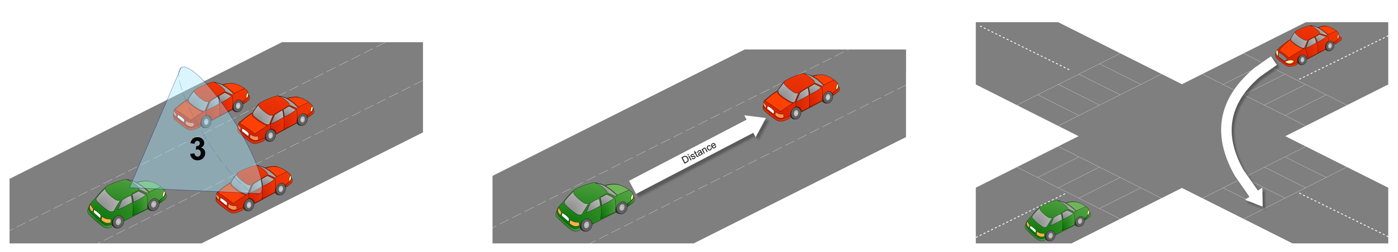
CVRR Computer Vision and Robotics Research Laboratory

LISA: Laboratory for Intelligent & Safe Automobiles

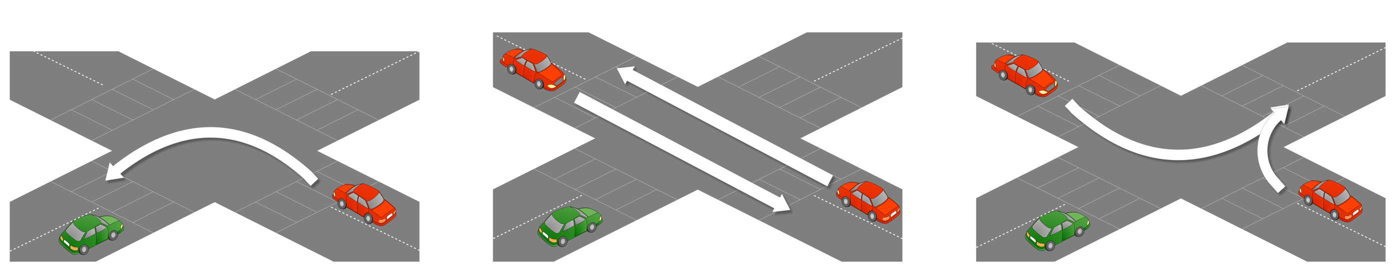
Introduction

We introduce a stereo-vision system for automatic Naturalistic Driving Study (NDS) data reduction on both day and nighttime data. The purpose of NDS is to provides insight in the patterns and behaviors of drivers during near-crashes and crashes. Since NDS in conducted on large amounts of data, it is extremely time consuming to analyses. Automated solutions, that can sort through and categorize the collected data, are therefore sought after.

A stereo-vision based system enables detection and tracking of vehicles in scenarios that would be problematic for monocular detectors. The system proposed in this paper will be limited to handling a handful of NDS events, which especially benefit from the extra dimension in stereo-vision. The NDS 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 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: Drive events that can be automatically detected by the method proposed in this paper. Green car is ego vehicle. Red cars are other vehicles.

Our contributions are:

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

References

[1] H. Hirschmuller, “Accurate and efficient stereo processing by semi-global matching and mutual information,” in *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, vol. 2, June 2005, pp. 807–814 vol. 2.

[2] R. Labayrade, D. Aubert, and J.-P. Tarel, “Real time obstacle detection in stereovision on non flat road geometry through" v-disparity" representation,” in *Intelligent Vehicle Symposium, 2002. IEEE*, vol. 2. IEEE, 2002, pp. 646–651.

[3] A. Geiger, J. Ziegler, and C. Stiller, “Stereoscan: Dense 3d reconstruction in real-time,” in *Intelligent Vehicles Symposium (IV)*, 2011.

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. An overview of the system is shown in Figure 2.

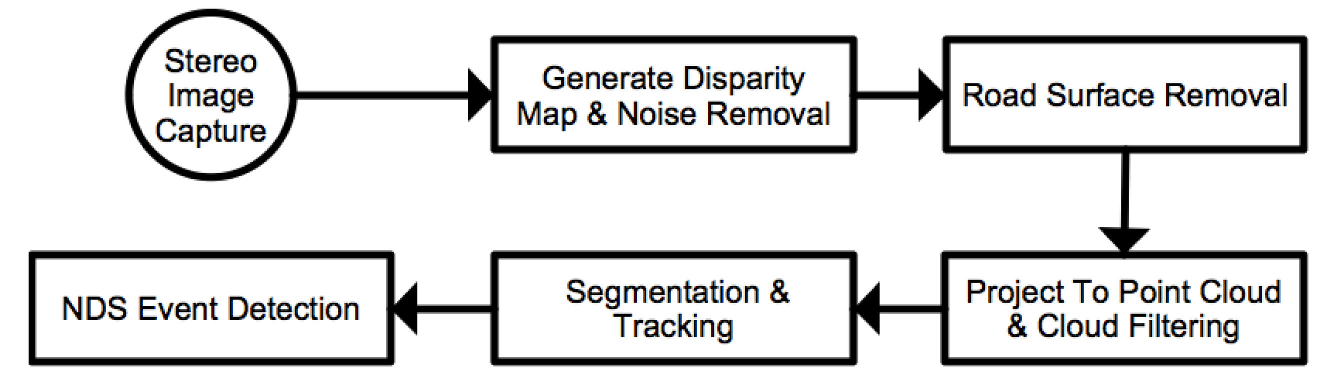


Figure 2: Overview of the methods used in the proposed NDS system.

Generate Disparity Map & Noise Removal

A disparity map is generated using OpenCV's SGBM implementation. For noise reduction, a LR-RL consistency check is done [1]. In case of a disparity differences in the consistency check, the lowest value is selected as the output pixel. A temporal consistency check assumes that a pixel's disparity value does not change erratically between frames. Finally, a monocular color check removes pixels from extremely light and dark regions.

Road Surface Removal

The road surface is found by searching for the most significant line in the V-disparity[2] using RANSAC. Additionally, the line parameters are filtered using a Kalman filter to smooth out faulty road surface detections. The parameters are then used to remove pixels for everything not above the road surface. The remaining pixels are considered vehicle candidate disparity pixels. In Figure 4 the left input image and the vehicle candidate disparity map are shown.

(a) A rectified image captured **(b)** Noise reduced disparity image.

Figure 3: Results after Generate Disparity Map & Noise Removal and Road Surface Removal.

Project To Point Cloud & Cloud Filtering

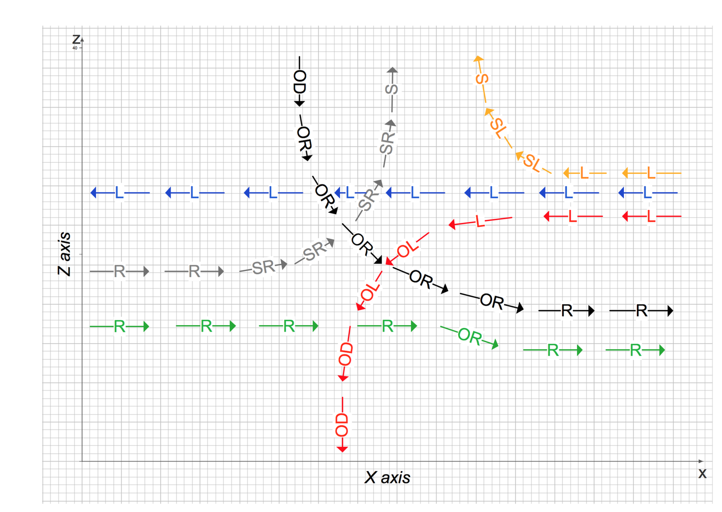
Using the camera's focal length f(in pixels) and it's baseline b(in meters), along with the calculated disparity d(in pixels), the actual distance z to objects in the camera's view can be found. The remaining x and y component of the world coordinate are found using the column and row index of the disparity pixel.

Segmentation & Tracking

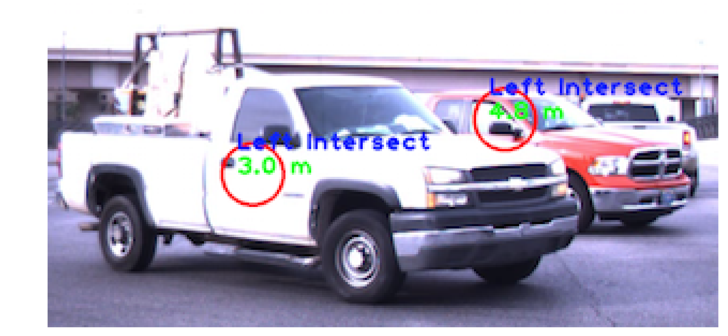
The acquired point clouds are post processed using a band-bass filter to remove near and distant points, downsampled using a voxel grid, and removing outliers. Clusters are found by creating k-d tree, and for each point finding neighbors within a specified radius. The clusters center point is used for nearest neighbor tracking between frame and used to determine the distance from ego vehicle to detected vehicle.

NDS Event Detection

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* [3]. NDS events are detected by looking at the movement history of other vehicles with regard to the ego vehicle.



(a)



(b)

Figure 4: (a) Illustration of various movements that are detected by the system. (b) NDS detection example.

For all detected vehicles, individual frame to frame movements are categorized to form a a histogram of events for determining which NDS events have occurred.

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 NDS events manually labeled and SO is the system output.

Table 1: Summary of drive analysis from NDS. 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

Conclusion

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. The proposed system have an overall precision of 0.9375 and an overall recall of 0.71, which is considered useful for data reduction in NDS. Future work may include looking at additional NDS events where stereo-vision can be utilized.

Acknowledgment

We would like to thank Ravi K. Satzoda, Mohan M. Trivedi, Andreas Møgelmose, and Thomas B. Moeslund for their guidance.