Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Institute for
Dynamic Systems and Control
Institut für Dynamische Systeme
und Regelungstechnik

Noah Isaak

Free Space Segmentation for Gokart Application

V-Disparity

Semester Project

Institute for Dynamic Systems and Control
Swiss Federal Institute of Technology (ETH) Zurich

Supervision

First Supervisor
Prof. Dr. Second Supervisor

December 2018

IDSC-XX-YY-ZZ

Abstract

Road and ground detection are closely related key tasks for an autonomous ground vehicle. These computations should be robust and preferably be performed in real-time. This paper aims to show the implementation of the V-Disparity method for ground detection. The approach is based on classic computer vision and does not incorporate learning methods. Basis of the method is a disparity map, for which a row-wise histogram is computed. This V-disparity histogram robustly preserves geometric scene features and can be used for various tasks. Experimental results however show the shortcomings of the implementation and how they could be overcome. In the following, a second pipeline is introduced. A mapping from a Velodyne Vld-16 Lidar to a camera is computed. The binary obstacle - ground mask which is computed from the lidar's point cloud can then be projected onto the camera image. These labels can then be used for e.g. machine learning tasks.

Keywords: Ground detection, V-Disparity.

Contents

Chapter 1

Introduction

The task of segmenting a scene into ground, obstacles and other labels is a well-researched topic. It is a fundamental part of any autonomous ground vehicle, and lays the basis for many different tasks, such as planning, safety features or scene understanding. Over the last few years, solutions to this task, which are based on machine learning methods, have deemed themselves to be robust and accurate. Their ability to generalise and be applied to a variety of scenes make them a reliable choice. There however, exist many approaches to the problem, which are based on classic computer vision, with implementations dating many years back.
The goal of this semester project is the implementation of a robust and accurate free space detection pipeline based on classic computer vision. The results of this pipeline can be used to e.g. label data for machine learning. If the pipeline does not perform as expected, a second pipeline is introduced, where the focus entirely lies on labeling data for machine learning tasks.

This semester project is part of a more extensive research into autonomous driving, based a self-driving Gokart. The Gokart is basis for a variety of research topics, such as planning, control systems and sensor fusion. The testing environment is a modular indoor Gokart track. The ground is flat, with no uphill or downhill sections. The track is not affected by weather conditions.

1.1 Related Work

Related work on this topic is extensive. After limiting my research to the method of V-Disparity, early implementations can be dated back to 2002, with Labayrade et al. [1] providing a robust and accurate method for road detection, even for non-flat geometries. A study on the U-V Disparity method can be found in [2], providing a real-time implementation of the stereovision based scene analysis. The method was improved and adapted since it's introduction, becoming more robust and computationally efficient, as shown in [3]. Aside from free space detection, U-V-Disparity can be used solely for obstacle and pedestrian detection, with U-V-Disparity acting as the underlying basis for a SVM Classifier, where the extraced ROIs are used for training. Iloie et al. implemented this framework in [4].

1.2 Structure of this report

This report is structured as follows. In 2 both, the V-Disparity method and Lidar to Camera projection is explained. Section 3 discusses the implementation of the above methods. Additionally, the mounted hardware is mentioned. Results are presented in section 4, followed by a conclusion and proposed future work in 5.

Chapter 2

Method

2.1 V-Disparity

This chapter explains the theoretical basis of the V-disparity method. To begin with, the basis of the V-Disparity method, the disparity map, is outlined.

2.1.1 Disparity Map

It is assumed, that the reader has a general understanding of stereovision systems. Given two calibrated stereo cameras, the respective rectified images can be used to calibrate a disparity for each pixel. These is achieved with different block-matching algorithms, such as the StereoSGBM algorithm implemented in the popular OpenCV library. Given a disparity map, the baseline and focal length, a depth map can be calculated and used for further processing. The V-Disparity method however only makes use of the disparity map.

2.1.2 V-Disparity

Once a disparity map δ is calculated, a V-Disparity histogram can be constructed. For each row u, a histogram of the occuring disparities is computed. The pixel values represent the occurence of the respective disparity in the row. Given a plane in the scene, the projection of the plane onto the V-Disparity image has a useful property. Following the calculations of 2.1.1, a plane will be projected as a linear curve in the V-Disparity image. This simplifies the extraction of the respective plane in the V-Disparity image, as a e.g. Hough Line Transform can be applied to detect the lines. Because the most prominent plane in usually represented by the ground, it will be detected as the line with the most votes in the V-Disparity image. Horizontal or vertical lines can be dismissed, as the disparity gradient of the road leads to a skewed line. Horizontal lines can be associated with obstacles and used for obstacle detection. Even though Non-flat road geometry will not be considered in this project, it should be noted that the road in that case can be approximated by a series of planes, which will then be projected as a piecewise linear curve in the V-Disparity image.

Once the line has been fitted in the V-Disparity image, the disparity values for the road surface are known. Extracting the road in the image domain is straightforward. For each row, the values which lie within a threshold of the value of the extracted line are part of the road, all other pixels are masked as non-road.

2.2 Lidar Camera Projection

To project lidar scanning points onto a camera image, the camera's intrinsic parameters should be known. Additionally, the projection matrix from lidar to camera should be known. It can be roughly measured and computed, or determined by externally calibrating the lidar and camera.

Chapter 3

Implementation

3.1 V-Disparity Implementation

3.2 Hardware

The Gokart is fitted with a ZED Stereolabs camera, which was the basis for the V-Disparity method. The camera comes with a ZED SDK, which provides many functionalities, such as disparity map generation and point cloud computation. In addition, a Velodyne VLD-16 Lidar is mounted on top of the Gokart. It offers 16 scanning lines with high accuracy. More infos on the used hardware can be found on the manufacturer's website.

3.3 Code implementation

The method was implemented in the Python framework, while making use of the popular OpenCV library. OpenCV was used, mainly because of the existing methods already implemented, such as the probabilistic houghline transform, which was used for line-fitting. The OpenCV aims at real-time computer vision programming functions, which is a crucial property for autonomous vehicles.

Given the ZED Camera and the ZED SDK, data can be streamed from the Camera with few lines of code. Depending on the application, the resolution, framerate and other parameters can be adjusted accordingly. For disparity computation, range and quality can be modified, to fit the requirements. The disparity map is computed as float32. For visualisation purposes, a normalisation and conversion to uint8 type needs to be done. For higher precision, the float32 disparity map can be used for further processing. To ease up on computation, the uint8 disparity map was used for the V-Disparity method. The number of histogram bins for every row histogram is then set to 255, one for every disparity value present in the image. OpenCV's calcHist function was used.

In order to avoid speckles and disjointed mask segments, all pixels are labeled according to their connectedness, using OpenCV's connectedComponents function. Only the largest connected region is kept, all other labels are discarded.

Chapter 4

Results

Results have been visually evaluated, as there is no ground truth available for this method. The evaluation have shown, that for this project's implementation, the V-Disparity method does not provide the desired results. The generated mask lacks temporal coherency, and for mid-range distances (>4m), important features are not detected, such as various obstacles or road boundaries. The reason for this can be found in the generated disparity map. Without post processing, the backprojection of the road requires an accurate disparity map. However the quality of the disparity map diminishes for further distances, edges are not well preserved and artifacts in low-textured areas lead to inaccurate masks. As the disparity values for low-textured areas are hard to determine via block matching and the underlying disparity computation code is closed source, the exact reason for the error is hard to determine.
Adjustments of the parameters can lead to better results. An extensive study on the quality and error of the ZED Camera was carried out by Ortiz et al. [5].

Chapter 5

Conclusion

Appendix A

AppendixChapter

Bibliography

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

[2] Z. Hu and K. Uchimura, “U-V-disparity: An efficient algorithm for stereovision based scene analysis,” *IEEE Intelligent Vehicles Symposium, Proceedings*, vol. 2005, pp. 48–54, 2005.

[3] S. Kakegawa, H. Matono, H. Kido, and T. Shima, “Road Surface Segmentation based on Vertically Local Disparity Histogram for Stereo Camera,” *International Journal of Intelligent Transportation Systems Research*, vol. 16, no. 2, pp. 90–97, 2018.

[4] A. Iloie, I. Giosan, and S. Nedevschi, “UV disparity based obstacle detection and pedestrian classification in urban traffic scenarios,” *Proceedings - 2014 IEEE 10th International Conference on Intelligent Computer Communication and Processing, ICCP 2014*, pp. 119–125, 2014.

[5] L. E. Ortiz, E. V. Cabrera, and L. M. Gonc, “Depth Data Error Modeling of the ZED 3D Vision Sensor from Stereolabs,” vol. 17, no. 1, pp. 1–15, 2018.

Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Institute for Dynamic Systems and Control

Prof. Dr. R. D'Andrea, Prof. Dr. E. Frazzoli, Prof. Dr. C. Onder, Prof. Dr. M. Zeilinger

Title of work:

Free Space Segmentation for Gokart Application

V-Disparity

Thesis type and date:

Semester Project, December 2018

Supervision:

First Supervisor
Prof. Dr. Second Supervisor

Student:

Name:	Noah Isaak
E-mail:	nisaak@student.ethz.ch
Legi-Nr.:	13-929-476
Semester:	FS 2019

Statement regarding plagiarism:

By signing this statement, I affirm that I have read and signed the Declaration of Originality, independently produced this paper, and adhered to the general practice of source citation in this subject-area.

Declaration of Originality:

`https://www.ethz.ch/content/dam/ethz/main/education/rechtliches-abschluesse/leistungskontrollen/declaration-originality.pdf`

Zurich, 28.5.2019: ____________________