Grid Map based Free Space Estimation using Stereo Vision

Hannes Harms[1], Eike Rehder[1] and Martin Lauer[1]

***Abstract*— This contribution proposes a temporally filtered free space estimation method for autonomous driving using dense disparity images from stereo vision. Urban environments feature complex surroundings in which the free space is limited by large and relatively flat obstacles (e.g. cars and curbs). Free space methods relying on single frame measurements suffer from sensor noise and depth artifacts, leading to large deviations from the ground truth free space. We meet this challenge by temporally filtering the occupancy of height and orientation features in a probabilistic occupancy grid. In the following the free space boundaries are estimated from the origin of the sensor in the occupancy grid. In contrast to existing methods our approach allows us to detect static free space boundaries which are not in the current sensor's field of view. The proposed approach is evaluated on different urban scenes and compared to a state of the art free space method.**

I. Method

The introduced free space estimation method is defined as the area in front of the car which is limited by large obstacles or curbs. The proposed system filters the obstacle and curb information temporally in a multi-layer occupancy grid as introduced in [3]. In the following, we will explain the processing pipeline step by step, as shown in figure 1.

A. Input Data

The proposed method uses a calibrated stereo camera setup as input which delivers two synchronized and rectified images at each time step. A state of the art stereo matcher [4] computes a dense *disparity image* for each image pair.

An *orientation image* is computed for each disparity image, which describes a pixel-wise deviation angle δ of a local normal vector w.r.t. a global height axis in world coordinates. The pixel-wise local normal vector is computed in disparity space as suggested in [5] and transformed into world coordinates. See [6] for details on the transformation. Note that δ is a dihedral angle, so that $\delta \in [0°, 90°]$. The global height axis $\mathbf{h}_{\text{world}}$ can be determined by an independent measurement from an IMU (acceleration and magnetic sensors) or estimated by tracking the ground plane normal vector over time. The orientation image I_δ contains the deviation angles δ for each pixel position (u,v).

The *motion* between two consecutive time steps is compensated by visual odometry [7].

B. Multi-Layer Grid Map

In contrast to existing approaches we represent the surrounding of the vehicle by a temporally filtered multi-layer occupancy grid map [3] where the grid corresponds to the ground level, which is obtained from extrinsic calibration. The coordinate frame of the grid remains fixed w.r.t. to the vehicle, i.e. the grid map moves with the vehicle while the vehicle orientation is stored in the grid. Thus the ego motion is compensated in a grid map only by shifting the grid accordingly and updating the orientation.

In our case we use two grid layers to represent two different properties of the environment, namely obstacles and curbs. Both grid layers are updated from stereo vision information. In stereo imaging, obstacles appear as a multitude of same depth values at the same azimuth which cover a relatively large height. We thus update the occupancy grid layers by column-wise processing of the image. Starting at the closest point of each viewing ray (=image column) relative to the camera, all cells of an occupancy grid layer are updated with a negative value in *logOdds* until an occupied cell is reached. The first occupied cell is updated with a positive value. The obstacle layer updates all grid cells until a cell with a height larger than 0.5 meters is detected. The curb layer updates all grid cells until the first obstacle cell or a cell with an orientation value δ larger than 30° is reached.

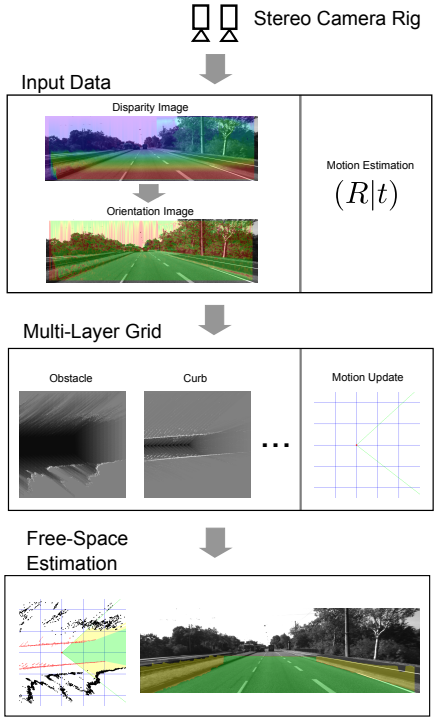

Fig. 1: Main steps of the curb and free space estimation pipeline using depth data only.

[1]Hannes Harms, Eike Rehder and Martin Lauer are with Insitute of Measurement and Control Systems, Karlsruhe Institute of Technology (KIT), 76131 Karlsruhe, Germany harms@kit.edu, eike.rehder@kit.edu, martin.lauer@kit.edu

(a) One lane street scene (b) Narrowing street (c) Expanding street (d) Wide street scene

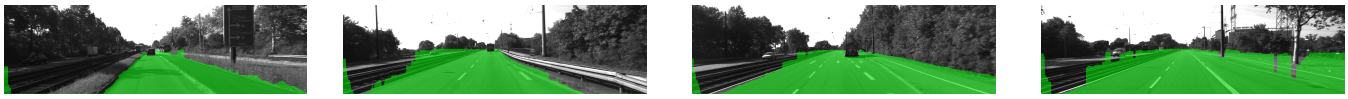

Fig. 2: Free space result images using a plane estimation as described in [1]

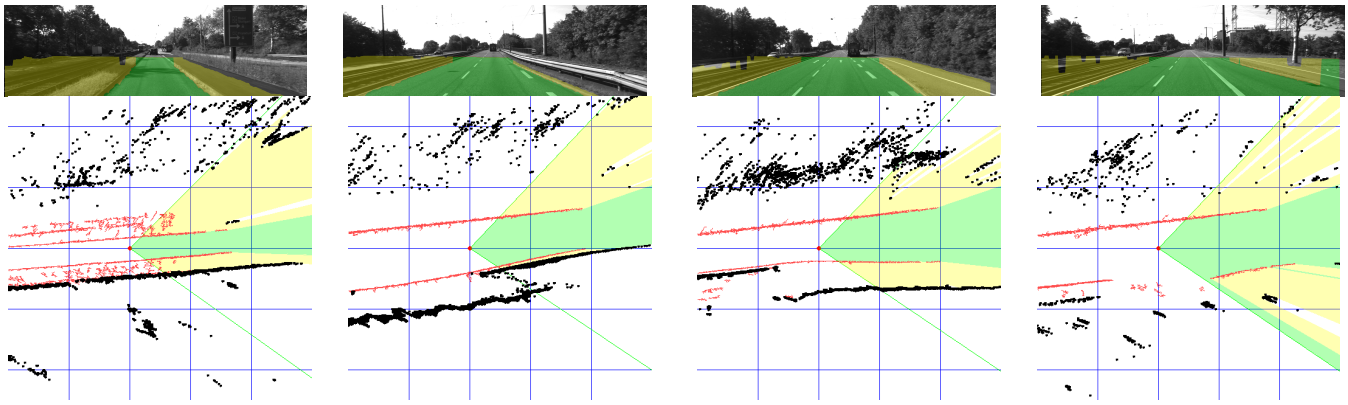

Fig. 3: Free space result images and corresponding occupancy grid representation from different urban scenarios. The green area represents the extracted free-space area limited by curbs, obstacles or a maximum distance while the yellow area shows the area behind the detected curb boundary. Obstacles are visualized as black circles and considered curb areas as red points.

C. Free Space Estimation

The proposed approach uses a ray based free space search in the occupancy grid. The maximum amount of rays and their directions are defined by the camera opening angle and horizontal resolution. The rays emanate from the camera origin until the first occupied cell in one of the occupancy grid layers is reached. Occupied cells are obtained by binarizing the obstacle and curb layer and applying morphological operations. The resulting free space is visualized in grid coordinates and projected into image space, see figure 3.

II. Results

The multi layer occupancy grid is initialized by a size of 50 x 50 m with a cell size of 0.1 x 0.1 m. The camera origin is positioned at $(20, 30)$ m and is marked as a thick red dot in the occupancy grid. The camera direction is signified in the occupancy grid by two green lines which mark the current field of view. Blue lines mark reference lines which intersect in distances of 10 m.

Our approach is designed to work on complex urban surroundings and long sequences, as can be found in the KITTI data set [8]. In comparison to the free space computed as described in [1], a significant refinement of the free space area can be achieved (see figure 2 and 3). Our proposed approach correctly estimates the free space on narrow and wide streets limited by curbs and large obstacles as well as on narrowing and expanding streets. Obstacles that limit the free space area like guard rails or vegetation are mapped and detected correctly.

On a 2.67 GHz dual core CPU the proposed method runs at approximately 10 Hz. Computational most expensive parts are stereo matching, normal computation, and odometry.

III. Conclusion and Outlook

In this contribution, we introduced a method for free space estimation which considers temporal filtered obstacles and curbs in a stochastic multi-layer occupancy grid as delimiters. First results show the benefits of the non-parametric estimation of obstacle and curb areas. The visibility of curbs reaches up to 30 meters, depending on the curb height and quality of depth information. The ray based free space estimation in the occupancy grid proved to be robust for different environments and has the advantage to be expandable by other sensors offering different fields of view. Since the system uses depth data only, other depth sensors (e.g. lidar) can substitute the used stereo camera rig.

References

[1] D. Pfeiffer and U. Franke, "Efficient representation of traffic scenes by means of dynamic stixels," in *Intelligent Vehicles Symposium (IV), 2010 IEEE*. IEEE, 2010, pp. 217–224.

[2] P. Y. Shinzato, D. Gomes, and D. F. Wolf, "Road estimation with sparse 3d points from stereo data," in *Intelligent Transportation Systems (ITSC)*. IEEE, 2014, pp. 1688–1693.

[3] A. Elfes, "Using occupancy grids for mobile robot perception and navigation," *Computer*, vol. 22, no. 6, pp. 46–57, 1989.

[4] B. Ranft and T. Strauß, "Modeling arbitrarily oriented slanted planes for efficient stereo vision based on block matching," in *Intelligent Transportation Systems (ITSC)*. IEEE, 2014, pp. 1941–1947.

[5] H. Badino, D. Huber, Y. Park, and T. Kanade, "Fast and accurate computation of surface normals from range images," in *Robotics and Automation (ICRA)*. IEEE, 2011, pp. 3084–3091.

[6] H. Harms, J. Beck, J. Ziegler, and C. Stiller, "Accuracy analysis of surface normal reconstruction in stereo vision," in *Intelligent Vehicles Symposium Proceedings, 2014 IEEE*. IEEE, 2014, pp. 730–736.

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

[8] A. Geiger, P. Lenz, C. Stiller, and R. Urtasun, "Vision meets robotics: The kitti dataset," *The International Journal of Robotics Research*, p. 0278364913491297, 2013.